

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 1999, 15:03:38 ; Search time 1096.08 Seconds
(without alignments)
35.993 Million cell updates/sec

Title: US-09-026-400-6

Perfect score: 20

Sequence: 1 gcdatrtgncraanacnc 20

Scoring table: IDENTITY_NUC

Searched: 2546578 seqs, 986266752 residues

Database :

EST:**

1: em_est1:**

2: em_est2:**

3: em_est3:**

4: em_est4:**

5: em_est5:**

6: em_est6:**

7: em_est7:**

8: em_est8:**

9: em_est9:**

10: em_est10:**

11: em_est11:**

12: em_est12:**

13: em_est13:**

14: em_est14:**

15: em_est15:**

16: em_est16:**

17: em_est17:**

18: em_est18:**

19: em_est19:**

20: gb_est1:**

21: gb_est2:**

22: gb_est3:**

23: gb_est4:**

24: gb_est5:**

25: gb_est6:**

26: gb_est7:**

27: gb_est8:**

28: gb_est9:**

29: gb_est10:**

30: gb_est11:**

31: gb_est12:**

32: gb_est13:**

33: gb_est14:**

34: gb_est15:**

35: gb_est16:**

36: gb_est17:**

37: gb_est18:**

38: gb_est19:**

39: gb_est20:**

40: gb_est21:**

41: gb_est22:**

42: gb_est23:**

43: gb_est24:**

44: gb_est25:**

45: gb_est26:**

46: gb_est27:**

47: gb_est28:**

48: gb_est29:**

49: gb_est30:**

50: gb_est31:**

51: gb_est32:**

52: em_est20:**

53: em_est21:**

54: em_est22:**
55: em_est23:**
56: em_est24:**
57: em_est25:**
58: em_est26:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	14.4	72.0	599	40	AA941765
2	14.4	72.0	785	40	AA952006
3	14.4	72.0	649	40	AA990815
4	14.4	72.0	652	41	AI061712
5	14.4	72.0	672	44	AI257428
6	14.4	72.0	649	44	AI259611
7	14.4	72.0	609	47	AI512710
8	14.4	72.0	613	47	AI512875
9	14.4	72.0	658	47	AI515803
10	14.4	72.0	586	47	AI531887
11	14.4	72.0	444	48	AI569754
12	13.8	69.0	643	33	AA390566
13	13.8	69.0	721	33	AA439525
14	13.8	69.0	645	33	AA440182
15	13.8	69.0	504	33	AA440803
16	13.8	69.0	414	37	AA580465
17	13.8	69.0	587	37	AA695021
18	13.8	69.0	720	37	AA704069
19	13.8	69.0	536	39	AA870850
20	13.8	69.0	326	39	AA883627
21	13.8	69.0	381	40	AA936014
22	13.8	69.0	657	40	AA941506
23	13.8	69.0	577	40	AA951871
24	13.8	69.0	332	40	AA969205
25	13.8	69.0	599	41	AI012608
26	13.8	69.0	411	43	AI176025
27	13.8	69.0	560	45	AI353745
28	13.8	69.0	439	46	AA926282
29	13.8	69.0	439	46	AA955061
30	13.8	69.0	615	47	AI514859
31	13.8	69.0	574	47	AI514937
32	13.8	69.0	581	47	AI516637
33	13.8	69.0	660	48	AI546199
34	13.8	69.0	654	48	AI546381
35	13.8	69.0	437	48	AI554779
36	13.8	69.0	546	48	AI571347
37	13.8	69.0	538	48	AI599587
38	13.8	69.0	504	50	AI684230
39	13.4	67.0	417	21	R11658
40	13.4	67.0	348	22	R80751
41	13.4	67.0	390	23	R95042
42	13.4	67.0	448	23	R95183
43	13.4	67.0	383	32	AA378645
44	13.4	67.0	438	33	AA411566
45	13.4	67.0	404	33	AA411567

ALIGNMENTS

RESULT 1

AA941765

LOCUS

DEFINITION

ACCESION

NID

VERSION

AA941765

LD25946.5prime LD Drosophila melanogaster embryo p0T2 Drosophila

melanogaster cDNA clone LD25946 5prime, mRNA sequence.

AA941765

93101598

599 bp

mRNA

EST

25-NOV-1998

Fri Oct 1 16:47:14 1999

KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 599)
Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
Brokstein, P., Lewis, S. and Rubin, G.M.
BDGP/HMI Drosophila EST Project
Unpublished (1997)
On Dec 18, 1997 this sequence version replaced gi:2339750.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>
Plate: 259 row: D column: 10
High quality sequence stop: 553.

FEATURES

SOURCE
1..599
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="LD25946"
/clone_lib="LD Drosophila melanogaster embryo pOT2"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="XL1 Blue"
/note="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2."
BASE COUNT 161 a 161 c 142 g 135 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 40; Length 599;
Best Local Similarity 68.4%; Pred. No. 1.5e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gcdatrtgncraanacnc 19
||:|||||:|||||
Db 543 GCGATGTGCCGGAACGC 561

RESULT 2
AA952006 785 bp mRNA EST 25-NOV-1998
LOCUS LD29007.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
DEFINITION melanogaster cDNA clone LD29007 5prime, mRNA sequence.
ACCESSION AA952006
NID 93111819
VERSION AA952006.1 GI:3111819
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 785)
Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
Brokstein, P., Lewis, S. and Rubin, G.M.
BDGP/HMI Drosophila EST Project
Unpublished (1997)
On Jan 19, 1998 this sequence version replaced gi:2287479.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>

Plate: 290 row: A column: 7
High quality sequence stop: 754.
FEATURES
SOURCE
Location/Qualifiers
1..785
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map="846F08"
/clone="LD29007"
/clone_lib="LD Drosophila melanogaster embryo pOT2"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="XL1 Blue"
/note="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2."
BASE COUNT 209 a 223 c 176 g 177 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 40; Length 785;
Best Local Similarity 68.4%; Pred. No. 1.6e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gcdatrtgncraanacnc 19
||:|||||:|||||
Db 554 GCGATGTGCCGGAACGC 572

RESULT 3
AA990815 649 bp mRNA EST 24-NOV-1998
LOCUS LD34717.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
DEFINITION melanogaster cDNA clone LD34717 5prime, mRNA sequence.
ACCESSION AA990815
NID 93177348
VERSION AA990815.1 GI:3177348
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 649)
Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
Brokstein, P., Lewis, S. and Rubin, G.M.
BDGP/HMI Drosophila EST Project
Unpublished (1997)
On Jan 19, 1998 this sequence version replaced gi:2287188.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: <http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu>
Plate: 347 row: B column: 5
High quality sequence stop: 529.

FEATURES
SOURCE

Location/Qualifiers
1..649
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="LD34717"
/clone_lib="LD Drosophila melanogaster embryo pOT2"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="XL1 Blue"
/note="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2."
BASE COUNT 173 a 175 c 158 g 143 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 40; Length 649;
 Best Local Similarity 68.4%; Pred. No. 1.5e+02;
 Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gcdatrtgncraanaacnc 19
 Db 521 GCGATGTGCCGGAACGC 539

RESULT 4
 AI061712 EST 24-NOV-1998
 LOCUS LD33793.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
 DEFINITION melanogaster cDNA clone LD33793 5prime, mRNA sequence.
 ACCESSION AI061712
 NID 93337628
 VERSION AI061712.1 GI:3337628
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 652)
 AUTHORS Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
 Brokstein, P., Lewis, S. and Rubin, G.M.
 TITLE BDGP/HMI Drosophila EST Project
 JOURNAL Unpublished (1997)
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2152369.

Contact: Harvey, D.
 G. M. Rubin-Molecular and Cell Biology
 University of California Berkeley
 539 LSA, Berkeley, CA 94720-3200, USA
 Fax: 510 643 9947
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 Plate: 337 row: H column: 9
 High quality sequence stop: 576.
 Location/Qualifiers

FEATURES

source
 1..652
 /db_xref="taxon:7227"
 /map="21q"
 /clone="LD33793"
 /clone_lib="LD Drosophila melanogaster embryo pOT2"
 /sex="male and female"
 /dev_stage="0 to 24 hours mixed stage embryonic"
 /lab_host="X11 Blue"
 /note="Organ: embryo; Vector: pOT2; Site_1: EcoRT; Site_2:
 XhoI; Sized fractionated cDNAs were directly ligated into
 pOT2."
 BASE COUNT 177 a 178 c 156 g 141 t
 ORIGIN

Query Match 72.0%; Score 14.4; DB 41; Length 652;
 Best Local Similarity 68.4%; Pred. No. 1.5e+02;
 Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gcdatrtgncraanaacnc 19
 Db 533 GCGATGTGCCGGAACGC 551

RESULT 5
 AI257428 EST 17-NOV-1998
 LOCUS LP05788.5prime LP Drosophila melanogaster larval-early pupal pOT2
 DEFINITION Drosophila melanogaster cDNA clone LP05788 5prime, mRNA sequence.
 ACCESSION AI257428
 NID 93864953
 VERSION AI257428.1 GI:3864953
 KEYWORDS EST.

SOURCE
 ORGANISM
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 672)
 AUTHORS Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
 Brokstein, P., Lewis, S. and Rubin, G.M.
 TITLE BDGP/HMI Drosophila EST Project
 JOURNAL Unpublished (1997)
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2150302.

Contact: Harvey, D.
 G. M. Rubin-Molecular and Cell Biology
 University of California Berkeley
 539 LSA, Berkeley, CA 94720-3200, USA
 Fax: 510 643 9947
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 Plate: 57 row: H column: 4
 High quality sequence stop: 640.
 Location/Qualifiers

FEATURES
 source
 1..672
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /map="6p21.3"
 /clone="LP05788"
 /clone_lib="LP Drosophila melanogaster larval-early pupal
 pOT2"
 /sex="male and female"
 /dev_stage="larvae-pupae"
 /lab_host="DH5-alpha"
 /note="Organ: whole body; Vector: pOT2; Site_1: EcoRI;
 Site_2: XhoI; Sized fractionated cDNAs were directly
 ligated into pOT2. Plasmid cDNA library."

BASE COUNT 174 a 184 c 163 g 151 t
 ORIGIN

Query Match 72.0%; Score 14.4; DB 44; Length 672;
 Best Local Similarity 68.4%; Pred. No. 1.5e+02;
 Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gcdatrtgncraanaacnc 19
 Db 515 GCGATGTGCCGGAACGC 533

RESULT 6
 AI259611 EST 17-NOV-1998
 LOCUS LP03065.5prime LP Drosophila melanogaster larval-early pupal pOT2
 DEFINITION Drosophila melanogaster cDNA clone LP03065 5prime, mRNA sequence.
 ACCESSION AI259611
 NID 93867136
 VERSION AI259611.1 GI:3867136
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 649)
 AUTHORS Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
 Brokstein, P., Lewis, S. and Rubin, G.M.
 TITLE BDGP/HMI Drosophila EST Project
 JOURNAL Unpublished (1997)
 COMMENT On Aug 21, 1998 this sequence version replaced.

Contact: Harvey, D.
 G. M. Rubin-Molecular and Cell Biology
 University of California Berkeley
 539 LSA, Berkeley, CA 94720-3200, USA

Fax: 510 643 9947
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Plate: 30 row: F column: 5
High quality sequence stop: 483.
Location/Qualifiers
1..649
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="LP03065"
/clone_lib="LP Drosophila melanogaster larval-early pupal pOT2"
/sex="male and female"
/dev_stage="larvae-pupae"
/lab_host="DHS-alpha"
/note="Organ: whole body; Vector: pOT2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."

BASE COUNT 177 a 177 c 153 g 142 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 44; Length 649;
Best Local Similarity 68.4%; Pred. No. 1.5e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 gdatrtgncraaacnc 19
||:||||| ||:|||||
Db 548 GCGATGTGCCGAAAACGC 566

RESULT 7
A1512710 609 bp mRNA EST 16-MAR-1999
LOCUS LD44673.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
DEFINITION melanogaster cDNA clone LD44673 5prime, mRNA sequence.
ACCESSION A1512710
NID 94422128
VERSION A1512710.1 GI:4422128
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 609)
AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
Brokstein,P., Lewis,S. and Rubin,G.M.
TITLE BDGP/HMMI Drosophila EST Project
JOURNAL Unpublished (1997)
COMMENT On May 18, 1998 this sequence version replaced gi:3136738.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Plate: 446 row: G column: 1
High quality sequence stop: 513.
Location/Qualifiers
1..609
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="LD44673"
/clone_lib="LD Drosophila melanogaster embryo pOT2"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="Xhl Blue"
/note="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2."

BASE COUNT 165 a 166 c 148 g 130 t

FEATURES
source

ORIGIN

Query Match 72.0%; Score 14.4; DB 47; Length 609;
Best Local Similarity 68.4%; Pred. No. 1.5e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 gdatrtgncraaacnc 19
||:||||| ||:|||||
Db 523 GCGATGTGCCGAAAACGC 541

RESULT 8
A1512875 613 bp mRNA EST 16-MAR-1999
LOCUS LD44958.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
DEFINITION melanogaster cDNA clone LD44958 5prime, mRNA sequence.
ACCESSION A1512875
NID 94422293
VERSION A1512875.1 GI:4422293
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 613)
AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
Brokstein,P., Lewis,S. and Rubin,G.M.
TITLE BDGP/HMMI Drosophila EST Project
JOURNAL Unpublished (1997)
COMMENT On May 18, 1998 this sequence version replaced gi:3136948.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Plate: 449 row: E column: 10
High quality sequence stop: 581.
Location/Qualifiers
1..613
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="LD44958"
/clone_lib="LD Drosophila melanogaster embryo pOT2"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="Xhl Blue"
/note="Organ: embryo; Vector: pOT2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2."

BASE COUNT 166 a 168 c 143 g 136 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 47; Length 613;
Best Local Similarity 68.4%; Pred. No. 1.5e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 gdatrtgncraaacnc 19
||:||||| ||:|||||
Db 554 GCGATGTGCCGAAAACGC 572

RESULT 9
A1515803 658 bp mRNA EST 16-MAR-1999
LOCUS LD44475.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
DEFINITION melanogaster cDNA clone LD44475 5prime, mRNA sequence.
ACCESSION A1515803
NID 94418903

```

VERSION      AI515803.1  GI:4418903
KEYWORDS     EST.
SOURCE       fruit fly.
ORGANISM     Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
             Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
             Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1 (bases 1 to 658)
AUTHORS      Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
             Brokstein,P., Lewis,S. and Rubin,G.M.
TITLE        BDGP/HMI Drosophila EST Project
JOURNAL      Unpublished (1997)
COMMENT      On Mar 10, 1998 this sequence version replaced gi:2948693.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 414 row: G column: 3
High quality sequence stop: 464.
Location/Qualifiers
1..658
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="LD41475"
/clone_lib="LD Drosophila melanogaster embryo pOT2"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="XLI Blue"
/notes="organ: embryo; Vector: pOT2; Site.1: EcoRI; Site.2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2."
BASE COUNT   172 a 178 c 162 g 146 t
ORIGIN

Query Match      72.0%; Score 14.4; DB 47; Length 658;
Best Local Similarity 68.4%; Pred. No. 1.5e-02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 gdatrtgncraaacnc 19
||:||||| ||:|||||
Db 509 GCGATGTCGCCGAAACGC 527

RESULT 10
AI531687
LOCUS      SD02921.5prime SD Drosophila melanogaster Schneider L2 cell culture
DEFINITION  pOT2 Drosophila melanogaster cDNA clone SD02921 5prime, mRNA
sequence.
ACCESSION  AI531687
NID        94445822
VERSION    AI531687.1  GI:4445822
KEYWORDS   EST.
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
             Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
             Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 586)
AUTHORS    Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
             Brokstein,P., Lewis,S. and Rubin,G.M.
TITLE      BDGP/HMI Drosophila EST Project
JOURNAL    Unpublished (1997)
COMMENT    On Mar 10, 1998 this sequence version replaced gi:2948379.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA

```

```

Fax: 510 643 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 29 row: B column: 9
High quality sequence stop: 477.
Location/Qualifiers
1..586
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map="11q13.1"
/clone="SD02921"
/clone_lib="SD Drosophila melanogaster Schneider L2 cell
culture pOT2"
/lab_host="DH5-alpha"
/notes="Vector: pOT2; Site.1: EcoRI; Site.2: XhoI; Sized
fractionated cDNAs were directly ligated into pOT2.
Plasmid cDNA library."
BASE COUNT   158 a 160 c 138 g 130 t
ORIGIN

Query Match      72.0%; Score 14.4; DB 47; Length 586;
Best Local Similarity 68.4%; Pred. No. 1.5e-02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 gcdatrtgncraaacnc 19
||:||||| ||:|||||
Db 543 GCGATGTCGCCGAAACGC 561

RESULT 11
AI569754
LOCUS      tN05011.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone
DEFINITION  mRNA sequence.
AI569754 444 bp mRNA EST 14-APR-1999
tN05011.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2166741 3',
mRNA sequence.
ACCESSION  AI569754
NID        94533128
VERSION    AI569754.1  GI:4533128
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 444)
AUTHORS    NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
             National Cancer Institute / National Institute of Neurological
             Disorders and Stroke, Brain Tumor Genome Anatomy Project
             (CGAP/BTGP), Tumor Gene Index
TITLE      Unpublished (1998)
JOURNAL    Unpublished (1998)
COMMENT    On Mar 10, 1998 this sequence version replaced gi:2948525.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 472 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 430.
Location/Qualifiers
1..444
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2166741"
/tissue_type="anaplastic oligodendroglioma"

```

/lab_host="DH10B"
 /note="Organ: brain; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAATGGAGCGCGCATAGGTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."
 BASE COUNT 97 a 129 c 130 g 87 t 1 others
 ORIGIN

Query Match 72.0%; Score 14.4; DB 48; Length 444;
 Best Local Similarity 68.4%; Pred. No. 1.4e+02;
 Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 cdatttgnccraanacnc 20
 ||:|||||:||||| ||
 Db 201 CAATGTGCGCGAAACACC 219

RESULT 12
 AA390566/c 643 bp mRNA EST 28-NOV-1998
 LOCUS LD09416.5prime LD Drosophila melanogaster embryo Bluescript
 DEFINITION Drosophila melanogaster cDNA clone LD09416 5prime, mRNA sequence.
 ACCESSION AA390566
 NID q2043560
 VERSION AA390566.1 GI:2043560
 KEYWORDS EST
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 643)
 AUTHORS Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
 Brokstein, P., Lewis, S. and Rubin, G.M.
 TITLE BDGP/HMI Drosophila EST Project
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1395289.

Contact: Harvey, D.
 G. M. Rubin-Molecular and Cell Biology
 University of California Berkeley
 539 LSA, Berkeley, CA 94720-3200, USA
 Fax: 510 643 9947
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 Plate: 94 row: B column: 4
 High quality sequence stop: 578.
 Location/Qualifiers
 1. 643
 /organism="Drosophila melanogaster"
 /db_xref="BDGP_EST:BDcln008715"
 /db_xref="taxon:7227"
 /clone="LD09416"
 /clone_lib="LD Drosophila melanogaster embryo Bluescript"
 /sex="male and female"
 /dev_stage="0 to 24 hours mixed stage embryonic"
 /lab_host="SOLR"
 /note="Organ: embryo; Vector: Bluescript SK; Site_1: EcoRI; Site_2: XhoI; Constructed using Stratagene ZAP-cDNA Synthesis kit. Oligo dT-primed and directionally cloned at EcoRI and XhoI in Bluescript SK(+/-)"

BASE COUNT 168 a 176 c 155 g 144 t
 ORIGIN

Query Match 69.0%; Score 13.8; DB 33; Length 643;
 Best Local Similarity 65.0%; Pred. No. 3.4e+02;
 Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 gdatrtgncraanacnc 20
 ||:|||||:||||| ||
 Db 260 GCAATGTGCGCGAAATCTCC 241

QY 1 gdatrtgncraanacnc 20
 ||:|||||:||||| ||
 Db 289 GCAATGTGCGCGAAATCTCC 270

RESULT 13
 AA439525/c 721 bp mRNA EST 28-NOV-1998
 LOCUS LD14791.5prime LD Drosophila melanogaster embryo Bluescript
 DEFINITION Drosophila melanogaster cDNA clone LD14791 5prime, mRNA sequence.
 ACCESSION AA439525
 NID g2151406
 VERSION AA439525.1 GI:2151406
 KEYWORDS EST
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 721)
 AUTHORS Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C.,
 Brokstein, P., Lewis, S. and Rubin, G.M.
 TITLE BDGP/HMI Drosophila EST Project
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1394108.

Contact: Harvey, D.
 G. M. Rubin-Molecular and Cell Biology
 University of California Berkeley
 539 LSA, Berkeley, CA 94720-3200, USA
 Fax: 510 643 9947
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 Plate: 141 row: B column: 4
 High quality sequence stop: 613.
 Location/Qualifiers
 1. 721
 /organism="Drosophila melanogaster"
 /db_xref="BDGP_EST:BDcln013323"
 /db_xref="taxon:7227"
 /clone="LD14116"
 /clone_lib="LD Drosophila melanogaster embryo Bluescript"
 /sex="male and female"
 /dev_stage="0 to 24 hours mixed stage embryonic"
 /lab_host="SOLR"
 /note="Organ: embryo; Vector: Bluescript SK; Site_1: EcoRI; Site_2: XhoI; Constructed using Stratagene ZAP-cDNA Synthesis kit. Oligo dT-primed and directionally cloned at EcoRI and XhoI in Bluescript SK(+/-)"

BASE COUNT 187 a 198 c 179 g 157 t
 ORIGIN

Query Match 69.0%; Score 13.8; DB 33; Length 721;
 Best Local Similarity 65.0%; Pred. No. 3.4e+02;
 Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 gdatrtgncraanacnc 20
 ||:|||||:||||| ||
 Db 260 GCAATGTGCGCGAAATCTCC 241

RESULT 14
 AA440182/c 645 bp mRNA EST 28-NOV-1998
 LOCUS LD14791.5prime LD Drosophila melanogaster embryo Bluescript
 DEFINITION Drosophila melanogaster cDNA clone LD14791 5prime, mRNA sequence.
 ACCESSION AA440182
 NID g2151920
 VERSION AA440182.1 GI:2151920
 KEYWORDS EST
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

AUTHORS 1 (bases 1 to 645) Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein,P., Lewis,S. and Rubin,G.M.

TITLE BDGP/HMI Drosophila EST Project

JOURNAL Unpublished (1997)

COMMENT On Sep 12, 1996 this sequence version replaced gi:1394984.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Plate: 147 row: H column: 7
High quality sequence stop: 472.

FEATURES
Source Location/Qualifiers
1..845
/organism="Drosophila melanogaster"
/db_xref="BDGP_EST:BDcln013974"
/db_xref="taxon:7227"
/clone="LD14791"
/clone_lib="LD Drosophila melanogaster embryo Bluescript"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="SOLR"
/note="Organ: embryo; Vector: Bluescript SK; Site:1;
ECORI; Site:2: XhoI; Constructed using Stratagene ZAP-cDNA
Synthesis kit. Oligo dt-primed and directionally cloned at
ECORI and XhoI in Bluescript SK(+/-)"
BASE COUNT 168 a 170 c 158 g 149 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 33; Length 645;
Best Local Similarity 65.0%; Pred. No. 3.4e+02;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 gcdatttgnccraanacncc 20
||:||||| ||:| | | |
Db 303 GCAATGTGCCGAATCTCC 284

RESULT 15
AA440803/c 504 bp mRNA EST 28-NOV-1998
LOCUS
DEFINITION LD15683.Sprime LD Drosophila melanogaster embryo Bluescript
Drosophila melanogaster cDNA clone LD15683 Spime, mRNA sequence.
AA440803
ACCESSION
NID g2152681
VERSION AA440803.1 GI:2152681
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 504) Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein,P., Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1394584.

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Plate: 156 row: G column: 11

FEATURES
Source Location/Qualifiers
1..504
/organism="Drosophila melanogaster"
/db_xref="BDGP_EST:BDcln014830"
/db_xref="taxon:7227"
/clone="LD15683"
/clone_lib="LD Drosophila melanogaster embryo Bluescript"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="SOLR"
/note="Organ: embryo; Vector: Bluescript SK; Site:1;
ECORI; Site:2: XhoI; Constructed using Stratagene ZAP-cDNA
Synthesis kit. Oligo dt-primed and directionally cloned at
ECORI and XhoI in Bluescript SK(+/-)"
BASE COUNT 129 a 128 c 124 g 123 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 33; Length 504;
Best Local Similarity 65.0%; Pred. No. 3.2e+02;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 gcdatttgnccraanacncc 20
||:||||| ||:| | | |
Db 296 GCAATGTGCCGAATCTCC 277

Search completed: October 1, 1999, 15:03:42
Job time: 4386 sec

us-09-026-400-6.rst

Fri Oct 1 16:47:14 1999

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 1999, 15:34:44 ; Search time 1619.02 Seconds
(without alignments)
39.287 Million cell updates/sec

Title: US-09-026-400-6

Perfect score: 20

Sequence: 1 gcdatrtgncraanacnc 20

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database :

GenEmbl.*

1: gb_bal.*

2: gb_ba2.*

3: gb_om.*

4: gb_ov.*

5: gb_pat.*

6: gb_ph.*

7: gb_pl1.*

8: gb_pl2.*

9: gb_pr1.*

10: gb_pr2.*

11: gb_pr3.*

12: gb_ro.*

13: gb_st.*

14: gb_sts.*

15: gb_sy.*

16: gb_un.*

17: gb_vi.*

18: em_fun.*

19: em_htg.*

20: em_hum1.*

21: em_hum2.*

22: em_in.*

23: em_om.*

24: em_or.*

25: em_ov.*

26: em_pat.*

27: em_ph.*

28: em_pl.*

29: em_ro.*

30: em_sts.*

31: em_sy.*

32: em_un.*

33: em_vi.*

34: gb_htg1.*

35: gb_htg2.*

36: gb_in1.*

37: gb_in2.*

38: em_bal.*

39: em_ba2.*

40: em_hum3.*

41: em_hum4.*

42: gb_pr4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
c 1	15.4	77.0	87210 7	AB011485 Arabidopsis

C	2	14.4	72.0	37304	1	MLCL373
C	3	14.4	72.0	4837	1	RSU1968
C	4	14.4	72.0	358	14	G01494
C	5	14.4	72.0	3214	17	HBV131133
C	6	14.4	72.0	110000	34	CEY38E10_0
C	7	14.4	72.0	110000	34	CEY57A10_3
C	8	14.4	72.0	125188	34	HSUJ447F3
C	9	14.4	72.0	163492	35	AC007241
C	10	14.4	72.0	193347	35	AC007742
C	11	14.4	72.0	14993	36	CEY37G9
C	12	14.4	72.0	3375	37	AF038726
C	13	13.8	69.0	801	1	D84200
C	14	13.8	69.0	17913	1	D90848
C	15	13.8	69.0	2664	1	HECUREASE
C	16	13.8	69.0	10036	2	AF000299
C	17	13.8	69.0	43075	6	U88974
C	18	13.8	69.0	4713	7	SCYKL197C
C	19	13.8	69.0	3479	7	YSCPAS1P
C	20	13.8	69.0	99689	8	AC004557
C	21	13.8	69.0	76133	10	HSU71148
C	22	13.8	69.0	151316	11	AC005039
C	23	13.8	69.0	112911	11	AC005865
C	24	13.8	69.0	165516	11	AC005912
C	25	13.8	69.0	35944	11	HSTCRB4H
C	26	13.8	69.0	921	11	HSU70137
C	27	13.8	69.0	216293	11	HTCRBCHR9
C	28	13.8	69.0	232650	11	U66061
C	29	13.8	69.0	1999	12	AF072865
C	30	13.8	69.0	195764	34	AC006483
C	31	13.8	69.0	187647	34	CEY16B4
C	32	13.8	69.0	123435	34	CEY61B8
C	33	13.8	69.0	142347	34	HS349A12
C	34	13.8	69.0	90106	35	AC007499
C	35	13.8	69.0	178624	35	AC007779
C	36	13.8	69.0	50924	35	AC007822
C	37	13.8	69.0	95262	36	AC005148
C	38	13.8	69.0	32258	36	CEY42D1
C	39	13.8	69.0	42829	36	CEY10G4
C	40	13.8	69.0	1689	36	DMU91634
C	41	13.8	69.0	994	36	DMU47881
C	42	13.8	69.0	1623	36	DMU49856
C	43	13.6	68.0	66194	35	AC007813
C	44	13.4	67.0	13699	1	AF000869
C	45	13.4	67.0	10513	1	AF023464

ALIGNMENTS

RESULT	1	AB011485/c	LOCUS	DEFINITION	AB011485	87210 bp	DNA	PLN	05-FEB-1999
LOCUS		AB011485	Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone: MXH1, complete sequence.						
ACCESSION		AB011485							
NID		G2924728							
VERSION		AB011485.1	GI:2924728						
SOURCE		HTG.							
ORGANISM		Arabidopsis thaliana							
		clone:MXH1.							
		Arabidopsis thaliana							
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.							
REFERENCE		1 (sites)							
AUTHORS		Nakamura, Y.							
TITLE		Structural Analysis of Arabidopsis thaliana Chromosome 5. V							
JOURNAL		Unpublished (1998)							
REFERENCE		2 (bases 1 to 87210)							
AUTHORS		Nakamura, Y.							
TITLE		Direct Submission							
JOURNAL		Submitted (02-MAR-1998) to the DDBJ/EMBL/GenBank databases.							


```

/note="LHII subunit beta"
/codon_start=1
/transl_table=11
/product="peripheral light harvesting complex subunit beta"
/protein_id="AAB59006.1"
/db_xref="PID:g1754677"
/db_xref="PID:g1754678"
/db_xref="GI:1754677"
/translation="MTDDMDKVPPTGLTLAEAEVHKOLIDGTRVFGAIALFAHFLAA
IATPWLG"
gene 2141..2332
/feature="pucA"
CDS 2141..2332
/feature="pucA"
/note="LHII subunit alpha precursor"
/codon_start=1
/transl_table=11
/product="peripheral light harvesting complex subunit
alpha precursor"
/protein_id="AAB59007.1"
/db_xref="PID:g1754678"
/db_xref="GI:1754678"
/translation="MNNAKMVLVVRPTVGIPFLVACAIASFLVHMLVLTGMWGDY
YSGSFEASLSVSNATLLS"
mat_peptide 2141..2293
/feature="pucA"
/note="11 amino acids are cleaved from the C-terminus of
LHII alpha subunit precursor to form mature protein"
/product="peripheral light harvesting complex subunit
alpha"
gene 2490..3854
/feature="pucC"
CDS 2490..3854
/feature="pucC"
/function="essential for assembly of the peripheral light
harvesting complex"
/codon_start=1
/transl_table=11
/product="putative regulatory protein PucC"
/protein_id="AAB59008.1"
/db_xref="PID:g1754679"
/db_xref="GI:1754679"
/translation="MNRSLKMAVNRATVGRFLPFAEASEDLPLSRLLRLSMPQVS
VGMAMVLVLTNRVNVIVLEVPASIVGINISLPLLPAPFRALIGFKSDTHKSLGWR
RVPIWKGTLQWGFIMPFAIVLSGOESAAGPEWIGILSAVSEFLVAGVHTV
QTVGLATDLAPREDQPNVGLMVLLGMIVSALLFGMLEDFYHAKLIKVQGA
AVATMVFVIALWKKMARQRARQRLGDPSPREANGLETGPNARLLWVILGL
TLGFLSLVLEPFQGVLDMSVAATKLTAAVAGTLYGFANASVLSRGYDPHAMA
GWAAGVGLPFAFIITFSATLQSEPVFVTMMAGFGAGLFSHTLTATMRSAKPAQVG
LALGAMGAVQATSAGAGIAGGVFDRVIRRSATSSARRPYSLSTRNRCWSRPSS
SATG"
terminator 762 a 1541 c 1658 g 876 t
BASE COUNT 3899..3935
ORIGIN

Query Match 72.0%; Score 14.4; DB 1; Length 4837;
Best Local Similarity 68.4%; Pred. No. 2.4e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 gcdatrtgncraanaacnc 19
||:|||||:|||||
Db 1106 CGCATATGACCGAATACTC 1088

RESULT 4
G01494 358 bp DNA STS 27-FEB-1995
LOCUS fruit fly STS Dm1847 clone DS07826 T7.
ACCESSION G01494
NID 9684897
VERSION G01494.1 GI:684897
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE fruit fly vector-PAdIOSacII The Pl library was made by D. Smoller

```

```

in D. Hartl's lab (see Smoller et al., Chromosoma 100: 487).
Chromosomal position was mapped by polytene chromosome in situ
hybridization in the Hartl lab (see Hartl et al., PNAS 91: 6824).
STS's were generated by sequencing the ends of the Drosophila
insert in these Pl clones, by the W. Kimmerly, C. Martin, and M.
Palazzolo lab at LBL.
Drosophila melanogaster
Eucaryotae; Metazoa; Arthropoda; Tracheata; Insecta; Diptera;
Brachycera; Cyclorhapha; Drosophilidae; Drosophila.
Rubin, G.
Drosophila STS
Unpublished (1994)

ORGANISM
REFERENCE 1 (bases 1 to 358)
AUTHORS Rubin, G.
TITLE Drosophila STS
JOURNAL Unpublished (1994)
COMMENT

Contact: Berkeley Drosophila Genome Project
Primer A: GATCTGTTTCATGTTGTCGG
Primer B: TTGGCCATATGCGTATATACA
STS size: 191
PCR Profile:
Annealing: 58 degrees C PCR Cycles: 32
Protocol:
Template: Pl Library Pools
Primer: 1 uM each
dNTPs: 250 uM each
Taq Poly: 0.05 units/ul
Total Vol: 15 ul
Buffer:
MgCl2: 1.5mM
KCl: 50 mM
Tris-HCl: 50 mM
pH: 8.3
Gelatin: .001 %

The Pl library has been distributed to 16 regional sites. A list
of these sites is available from FlyBase, via anonymous ftp to
ftp.bio.indiana.edu in the file
flybase/allied-data/genome-projects/1bl/LBL.doc.

FEATURES
source 1..358
/organism="Drosophila melanogaster"
STS 4..3358
primer_bind 4..24
BASE COUNT 110 a 75 c 67 g 102 t 4 others
ORIGIN

Query Match 72.0%; Score 14.4; DB 14; Length 358;
Best Local Similarity 68.4%; Pred. No. 2.2e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 gcdatrtgncraanaacnc 19
||:|||||:|||||
Db 182 GCAATATGCCAATAAC 200

RESULT 5
HBV131133/c 3214 bp DNA VRL 03-MAR-1999
LOCUS Hepatitis B virus, complete genome, strain J11.
DEFINITION AJ131133
ACCESSION 94468847
NID 94468847
VERSION AJ131133.1 GI:4468847
KEYWORDS C gene; core protein; P gene; polymerase; S gene; surface protein;
X gene; X protein.
SOURCE Hepatitis B virus.
ORGANISM Hepatitis B virus.
REFERENCE 1 (bases 1 to 3214)
AUTHORS Lu, L.C.
TITLE Direct Submission

```

JOURNAL Submitted (27-NOV-1998) Lu L.C., Microbiology and Immunology,
Leicester University, University Road, Leicester, LE1 9HN, UK

REFERENCE 2 (bases 1 to 3214)

AUTHORS Fujiyama, A., Miyanochara, A., Nozaki, C. and Yoneyama, T.

TITLE Point mutation in the S gene of hepatitis B virus for a d/y or w/r
subtypic change in two blood donors carrying a surface antigen of
compound subtype ady/r or advr

JOURNAL Unpublished

FEATURES

source Location/Qualifiers

1..3214 /organism="Hepatitis B virus"

/viation /strain="JUL1"

/db_xref="taxon:10407"

/lab_host="human hepatocytes"

/note="Hepadnaviridae"

1..3214 /gene="p"

1..3214 /gene="S"

1373..1837 /gene="X"

1373..1837 /gene="X"

/codon_start=1

/product="X protein"

/protein_id="CAB38227.1"

/db_xref="PID:el391948"

/db_xref="PID:94468848"

/db_xref="GI:4468848"

/translation="MDARLCQIDPARDVLCIRPVGAEGRGPFPGALPASPDPD
VPPDHGALRGLPVCFAFSAGCALRFTSARMTTVNAHNLKVLKHKRTGLSA
MSTDLKLVKDFNCEWELGEEVRLKVFVGLGCRHKLVCSAPCNFTSA"

1813..2451 /gene="C"

1813..2451 /gene="C"

/codon_start=1

/product="core protein"

/protein_id="CAB38228.1"

/db_xref="PID:el391949"

/db_xref="PID:94468849"

/db_xref="GI:4468849"

/translation="MQLFHLCLIIISCTPQVQAKLGLWGMQIDIPYKFGASVEL
LGFPSDFPISIRLDLTASLYREALSEPHSPHTALRQAILCWGELMNLATWVG
NLDEPAGRELVSVMNMGKLRQLLWFIISCLTFGRETVEYLVSGVWIRTTPA
YRPPNAPILSLPETTVRRGRSPRRTPSPRRRSQSPRRRSQSRSSQC"

Join(2305..3214,1..1284)

/gene="S"

/note="read-through stop codon"

/codon_start=1

/transl_except="(pos:3083..3085,aa:OTHER)

/product="polymerase"

/protein_id="CAB38229.1"

/db_xref="PID:el391950"

/db_xref="PID:94468850"

/db_xref="GI:4468850"

/translation="MPLSYQHFRKLLEDEAGLEELPRLADEGLNRVRAEDNLNG
NLNLSIPHTKVGNTFYSTVFCFNPKWOTSPFDIHLQEDIVDRCKQFVGLPTVN
ENRRLMPARFYPNVTKYLPDLKGIKPYPEHVNYFOTRHLTLWKAGILYKR
ESTRASFCGSPYNEODLHGRVLFOTSKRHGDKSPCPOSGILPRSSVGCIOSQL
RQSLGPOGAGOLAGROGGSIRARVHPFMGTGVPEXSRHTHCASFSSGF
NHRQVKAAVLSITSEGHSSGHAVELHFFPNSSRSQSGPVLFCWLOFRNSPFC
SQSGNTRVSWPKFAVNLQNLNLSLSLSDVSAAYFLPLHPAAPHLLVVG
SYLCUHVNIIDWGPCTEHEHRIKTPPTPVVGGVFLVDKNPHNTTESLVVDF
SSGLRYVARSLSNRSIINQHRTMNLNSCSRLNLYMLLYKTYGRKHLIYSHPI
ILGFRKIPMGVGLSPFLAQFTSAICSVRRAPPCCLAFSMDVVLGAKSVGHLES
YAAVTHPLGLIHLNPKYTKRWGYSNMGVYIGLGTIPQDHIVQKIKMWPKLPV
NRPIDMKVQRIVGLGFAEPTCGVPAALMYACIOAKAETSEPTYKAFILTKYL
NLPVVARPGCQVFADPTFGWGLAIGHQRMKGPCLLCRSILRNS"

Join(2712..3214,1..835)

/gene="S"

/codon_start=1

/product="surface protein"

/protein_id="CAB38230.1"

/db_xref="PID:el391951"

/db_xref="PID:94468851"

/db_xref="GI:4468851"

/translation="MMLIITSRDDIYTLFGRVSYIKESPRVAFPHFAGHILGNKIY
SMGWSKPRKMGNTLSVNPGLPFFDQDPAFAKANDNDPDLNPKHKNWPDANK
VGVGAFPGFIPSHGLGWSKAQOILTPAAASLASTIGKSRQPTPLSPPLURDT
HPQAMQWNTTFTQTDPRVRLYFPAGSSGTVSPAQNTVSAISSILSKTGDVVP
NMENIASGLGLPVLQAGFFLITKILITPQSLDSWWTSLNFGTPTVCLGNSQSQI
SSHSPCCPPICPGYRMCLRRFIFLCILLLCLIFLLVLDYQGMPLPCPLIPGSS
TSTGPKCTCTPAQGTSMFTSCCTKPTDGNCTCIPSSWAFAYLWEMASVRESWL
SLLVFFQMFVGLSPVWLSVIMMWYGFSLYNILSPFNPPLPIPFCLWYII"

BASE COUNT 734 a 842 c 728 g 910 t

ORIGIN

Query Match 72.0% Score 14.4; DB 17; Length 3214;
Best Local Similarity 68.4% Pred. No. 2.4e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 godatrtgncraanaanc 19
||:||||:||||:||||:
Db 910 GCAATGTGCCCAACACC 892

RESULT 6
CEY38E10_0/c
WPCOMMENT

Sequence split into 4 fragments LOCUS CEY38E10 Accession AL021149

Fragment Name	Begin	End
CEY38E10_0	1	110000
CEY38E10_1	100001	210000
CEY38E10_2	200001	310000
CEY38E10_3	300001	393707

LOCUS CEY38E10 393707 bp DNA HTG 21-MAY-1999

DEFINITION Caenorhabditis elegans chromosome II clone Y38E10, WORKING DRAFT
SEQUENCE, in unordered pieces.

ACCESSION AL021149

NID G4760249

VERSION AL021149.3 GI:4760249

KEYWORDS HTG; HTGS_PHASE1.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans

REFERENCE 1 (bases 1 to 393707)
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

AUTHORS Wallis, J.

TITLE Direct Submission

JOURNAL Submitted (21-MAY-1999) Nematode Sequencing Project, Sanger Centre,
Hinxton, Cambridge CB10 1RQ, UK and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
jes@sanger.ac.uk or rwenematode.wustl.edu

COMMENT On May 7, 1999 this sequence version replaced gi:4455375.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in progress
and the release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage etc. Order of segments is not known; 800 n's separate
segments.
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES

source Location/Qualifiers

1..393707 /organism="Caenorhabditis elegans"

/db_xref="taxon:6239"

/chromosome="II"

/clone="Y38E10"

BASE COUNT 117486 a 69904 c 68898 g 121367 t 16052 others

ORIGIN

Fri Oct 1 16:47:14 1999

```

Query Match      72.0%; Score 14.4; DB 34; Length 110000;
Best Local Similarity 68.4%; Pred. No. 2.8e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gcdatrtgncraaanacc 19
||:|||||:|||||
Db 24821 GCTATATGCCCCAAAAGCC 24803

RESULT 7
CEY57A10_3
WPCOMMENT
Sequence split into 5 fragments LOCUS CEY57A10 Accession AL020986
Fragment Name      Begin      End
CEY57A10_0         1      110000
CEY57A10_1        100001    210000
CEY57A10_2        200001    310000
CEY57A10_3        300001    410000
CEY57A10_4        400001    469134
Continuation (4 of 5) of CEY57A10 from base 300001 (AL020986 Caenorhabditis elegans chr3)

Query Match      72.0%; Score 14.4; DB 34; Length 110000;
Best Local Similarity 68.4%; Pred. No. 2.8e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gcdatrtgncraaanacc 19
||:|||||:|||||
Db 34404 GCTATATGCCCCAAAAGCC 34422

RESULT 8
HSDJ447F3
LOCUS
DEFINITION Homo sapiens chromosome 20 clone DJ447F3, WORKING DRAFT SEQUENCE,
in unordered pieces.
ACCESSION AL050348
NID 95051848
VERSION AL050348.6 GI:5051848
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 125188)
AUTHORS Lloyd.D.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-1999) Wellcome Trust Genome Campus, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jun 12, 1999 this sequence version replaced gi:5051342.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in progress
and the release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E. coli, yeast, vector,
phage etc. Order of segments is not known; 800 n's separate
segments. Unfinished: dj447F3 Contig_ID: 02039 acc=AL050348
Length: 23673 bp Unfinished: dj447F3 Contig_ID: 00738
acc=AL050348 Length: 8518 bp Unfinished: dj447F3 Contig_ID:
acc=AL050348 Length: 1653 bp Unfinished: dj447F3 Contig_ID:
01398 acc=AL050348 Length: 9114 bp Unfinished: dj447F3
Contig_ID: 01194 acc=AL050348 Length: 75081 bp Unfinished:
dj447F3 Contig_ID: 00234 acc=AL050348 Length: 1193 bp
Unfinished: dj447F3 Contig_ID: 02054 acc=AL050348 Length: 1156
bp.
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
Location/Qualifiers
FEATURES --
source
1..125188
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/clone="DJ447F3"
BASE COUNT 32970 a 29506 c 26656 g 31253 t 4803 others
ORIGIN
Query Match. 72.0%; Score 14.4; DB 34; Length 125188;
Best Local Similarity 68.4%; Pred. No. 2.8e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 cdatrtnccraaanacc 20
||:|||||:|||||
Db 122355 CTATGTGTCACAAATACCCC 122373

RESULT 9
AC007241
LOCUS
DEFINITION Homo sapiens clone NH0157F24, WORKING DRAFT SEQUENCE, 10 unordered
pieces.
ACCESSION AC007241
NID 94572725
VERSION AC007241.1 GI:4572725
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 163492)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163492)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (07-APR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2104: contig of 2104 bp in length
* 2105: gap of unknown length
* 2123: contig of 2829 bp in length
* 4952: gap of unknown length
* 4970: contig of 4568 bp in length
* 9537: gap of unknown length
* 9538: contig of 4880 bp in length
* 14435: gap of unknown length
* 14436: contig of 7670 bp in length
* 22123: contig of 7670 bp in length
* 22124: gap of unknown length
* 22141: contig of 8202 bp in length
* 30343: gap of unknown length
* 30361: contig of 7803 bp in length
* 30362: gap of unknown length
* 38165: contig of 19924 bp in length
* 38183: gap of unknown length
* 58107: contig of 51618 bp in length
* 58125: gap of unknown length
* 109743: gap of unknown length
* 109761: contig of 53732 bp in length.
Location/Qualifiers
FEATURES
source
1..163492
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NH0157F24"

```

```

BASE COUNT      46854 a 34129 c 35757 g 46590 t 162 others
ORIGIN

Query Match      72.0%; Score 14.4; DB 35; Length 163492;
Best Local Similarity 68.4%; Pred. No. 2.8e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 gcdatrtgncraanacnc 19
||:||||| ||:|||||
Db 57501 GCGATATGTCACACAC 57519

RESULT 10
AC007742      193347 bp      DNA      HTG      05-JUN-1999
LOCUS      Homo sapiens clone NH0357E16, WORKING DRAFT SEQUENCE, 15 unordered
DEFINITION      pieces.
ACCESSION      AC007742
NID      95001504
VERSION      AC007742.1 GI:5001504
KEYWORDS      HTG; HTGS_PHASE1.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 193347)
AUTHORS      Waterston, R.H.
TITLE      The sequence of Homo sapiens clone
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 193347)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (05-JUN-1999) Genome Sequencing Center, Washington
MO 63108, USA
COMMENT      * NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      2216: contig of 2216 bp in length
*      2217      2234: gap of unknown length
*      2235      7707: contig of 5473 bp in length
*      7708      7725: gap of unknown length
*      7726      16249: contig of 8524 bp in length
*      16250      16267: gap of unknown length
*      16268      26418: contig of 10151 bp in length
*      26419      26436: gap of unknown length
*      35911: contig of 9475 bp in length
*      35929: gap of unknown length
*      35930      45721: contig of 9792 bp in length
*      45722      45739: gap of unknown length
*      45740      55585: contig of 9846 bp in length
*      55586      55603: gap of unknown length
*      55604      66636: contig of 11033 bp in length
*      66637      66654: gap of unknown length
*      66655      78190: contig of 11536 bp in length
*      78191      78208: gap of unknown length
*      78209      90637: contig of 12429 bp in length
*      90638      90655: gap of unknown length
*      90656      104304: contig of 13649 bp in length
*      104305      104322: gap of unknown length
*      121503: contig of 17187 bp in length
*      121510      121527: gap of unknown length
*      121528      137176: contig of 15849 bp in length
*      137177      137193: gap of unknown length
*      160149: contig of 22956 bp in length
*      160150      160166: gap of unknown length
*      160167      193347: contig of 33181 bp in length.

FEATURES
source
Location/Qualifiers
1..193347
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NH0357E16"
BASE COUNT      58295 a 37110 c 37135 g 60557 t 250 others
ORIGIN

Query Match      72.0%; Score 14.4; DB 35; Length 193347;
Best Local Similarity 68.4%; Pred. No. 2.8e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 gcdatrtgncraanacnc 19
||:||||| ||:|||||
Db 121069 GCGATATGTCACACAC 121087

RESULT 11
CEF57G9      14993 bp      DNA      INV      23-NOV-1998
LOCUS      Caenorhabditis elegans cosmid F57G9, complete sequence.
DEFINITION      Z83231
ACCESSION      Z83231
NID      G1729610
VERSION      Z83231.1 GI:1729610
KEYWORDS      HTG
SOURCE      Caenorhabditis elegans.
ORGANISM      Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE      1 (bases 1 to 14993)
AUTHORS      Steward, C.
TITLE      Direct Submission
JOURNAL      Submitted (09-DEC-1996) Louis, MO 63110, USA. E-mail:
jes@sanger.ac.uk or rwenematode.wustl.edu
REFERENCE      2 (bases 1 to 14993)
AUTHORS      Wilson, R., Alnsough, R., Anderson, K., Baynes, C., Berks, M.,
Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,
Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,
Johnston, L., Jones, M., Kersey, J., Kirsten, J., Laister, N.,
Latreille, P., Lighthouse, J., Lloyd, C., McMurray, A., Mortimore, B.,
O'Callaghan, M., Parsons, J., Percy, C., Rifkin, L., Roopra, A.,
Saunders, D., Showkeen, R., Smalton, N., Smith, A., Sonhammer, E.,
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,
Wilkinson-Sproat, J. and Wohlman, P.
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans
Nature 368 (6466), 32-38 (1994)
JOURNAL      94150718
MEDLINE
COMMENT      Coding sequences below are predicted from computer analysis, using
predictions from Genefinder (P. Green, U. Washington), and other
available information.
For a graphical representation of this sequence and its analysis
see:-
http://webc.sanger.ac.uk/cgi-
bin/display?db=wormacsclass=Sequence object=F57G9
Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
IMPORTANT: This sequence is NOT necessarily the entire insert of
the specified clone. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
IMPORTANT: This sequence is not the entire insert of clone F57G9.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.
The true left end of clone F57G9 is at 41466 in
sequence AL032647.

```

The start of this sequence (1. .100) overlaps with the end of sequence AL023647.
The end of this sequence (14893. .14993) overlaps with the start of sequence AL023847.

FEATURES

Location/Qualifiers
1. .14993
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/chromosome="II"
/clone="F57G9"
/gene="F57G9.4"
complement(2578. .5070)
complement(join(2578. .2785,3390. .3514,3672. .3794,3836. .3992,4613. .5070))
/gene="F57G9.4"
/codon_start=1
/protein_id="CAB05752.1"
/db_xref="PID:el347402"
/db_xref="PID:g3877802"
/db_xref="GI:3877802"
/db_xref="SPTREMBL:O62269"
/translation="MLIKNSNDIWLPIHFYDETENFQVLVLSIVELFSLICAYILTL
NYIILKIMFERNYILAIFLQWHLIIGKLIITAYRLKILNPGEELGVHIEWT
SDPTRLKVESVNGLELLIFGGFLQWHYMTIFVGLAIVAVRVAIVLIEYNTQ
LPIPLFTVISTSLTALLFOKVGPLAQLPWIIICPFSAWAYFVFKCNESFP
REIRPRRHFVSQOQOVKENRALXIGTRLVFVLSIALCGIGITALEYDLIIP
FCHFVNEFLPHYLSCLTAFSVPQWNEFREYSLGRCLKIGRLKTESNAMEIQ
DSTKMGTTEDLYFQQLAUSWI"
complement(8130. .9673)
/gene="F57G9.2"
complement(join(8130. .8334,8393. .8517,8566. .8688,8745. .8901,9210. .9673))
/gene="F57G9.2"
/codon_start=1
/protein_id="CAB05750.1"
/db_xref="PID:el347400"
/db_xref="PID:g3877800"
/db_xref="GI:3877800"
/db_xref="SPTREMBL:O62267"
/translation="MLIHNSNSYIWLVSFVYKPLSLKLISIFELSSCILCGVIL
NLSIFVLKIQLFHKLMLTVPFAIWHLELIIGKTIITAYRLKIVNPGELGETV
WINDPKTLEVAGSSGLELLIFGGFLQWHYMTIFVGLAIVAVRVAIVLIEYNTQ
ERIPIITLISQLLSIAITQSGPLAFPLFVFCAPLSVLVLFKHTNQS
LKEICNPKRTFTVSQOQOVKENRALXIGTRLVFVLSIALCGIGITALEYDLIIP
PAGFGLIENFLPHYPLCTAFMFSIPQWQDFKKSILPFLNRRLAKIEQVVTVRIE
VYQNSSVETDIYERQLTESWT"
complement(10107. .11996)
/gene="F57G9.1"
complement(join(10107. .10311,10730. .10854,10907. .11029,11297. .11453,11533. .11996))
/gene="F57G9.1"
/codon_start=1
/protein_id="CAB05749.1"
/db_xref="PID:el347399"
/db_xref="PID:g3877799"
/db_xref="GI:3877799"
/db_xref="SPTREMBL:O62266"
/translation="MIKNGTGTFWLVPVYFNEPLNKLIVISIFELLSYILCGVIL
NLSYVMSKIQLEHKLMLTVPFAIWHLELIIGKTIITAYRLKIVNPGELGETV
WINDPKILVGGSSGLELLIFGGFLQWHYMTIFVGLAIVAVRVAIVLIEYNTQ
KRIWPIPIICOVLAIFMTFVINKVHPILARPLFCLPISFAVWLFPVKNKT
LQREIQNPKRTFTVSQOQOVKENRALXIGTRLVFVLSIALCGIGITALEYDLIIP
PGVCGHVENFLFHPICLTAFMFSIPQWQDFKKSILPFLNRRLAKIEQVVTVRIE
EDSKITSLFTDLYFKQLAESWI"
complement(12571. .13368)
/gene="F57G9.3"
complement(join(12571. .12996,13051. .13368))
/gene="F57G9.3"
/note="cDNA EST EMBL:T01304 comes from this gene"
/codon_start=1
/protein_id="CAB05751.1"
/db_xref="PID:el347401"
/db_xref="PID:g3877801"

BASE COUNT 5352 a 2542 c 2536 g 4563 t
ORIGIN
Query Match 72.0%; Score 14.4; DB 36; Length 14993;
Best Local Similarity 68.4%; Pred. No. 2.5e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 gdatatrgccraaacnc 19
||:||||:||||:||||:
Db 3094 GCTATATGCCAAACCC 3112

RESULT 12
AF028726/c
LOCUS
DEFINITION
Trypanosoma brucei kinetoplastid membrane proteins 1-4 (KMP-11/1, KMP-11/2, KMP-11/3, KMP-11/4) genes, nuclear genes encoding kinetoplastid proteins, complete cds.
ACCESSION
AF028726
NID
G3057129
VERSION
AF028726.1 GI:3057129
KEYWORDS
Trypanosoma brucei.
SOURCE
Trypanosoma brucei.
ORGANISM
Trypanosoma brucei.
Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
REFERENCE
1 (bases 1 to 3375)
Bridge,M.A., Zhou,Q., Koop,B.F. and Pearson,T.W.
Cloning and characterization of the kinetoplastid membrane protein-11 gene locus of Trypanosoma brucei
Mol. Biochem. Parasitol. 91 (2), 359-363 (1998)
JOURNAL
98226179
MEDLINE
2 (bases 1 to 3375)
Bridge,M.A., Zhou,Q., Koop,B.F. and Pearson,T.W.
Direct Submission
AUTHORS
Submitted (05-OCT-1997) Biochemistry and Microbiology, University of Victoria, P.O. Box 3055, Victoria, BC V8W 3P6, Canada
JOURNAL
Location/Qualifiers
FEATURES
source
1. .3375
/organism="trypanosoma brucei"
/db_xref="taxon:5691"
705. .983
/gene="KMP-11/1"
705. .983
/gene="KMP-11/1"
/note="membrane associated protein"
/codon_start=1
/product="kinetoplastid membrane protein 1"
/protein_id="AAC38990.1"
/db_xref="PID:g3057130"

/db_xref="GI:3877801"
/db_xref="SPTREMBL:O62268"
/translation="MGVGRILSVIYVFLCTLIACSVANNITVFGNDTAQVASCIV
VIITIFLILCSPICRTPREQLSIQFHINOQFIATSVIANVLVPLVMLGPK
NVSSIIICFLYCLLDOLGFAVEMIHFAFPIOTFPPKIKLSNPKYITILAESVIM
EFYHFCWNNISSNLPINILLCITAINIMEIPALIVAFYAINSDRTHFSRVKPGVGL
ELIIRKWDTKARTWKINESPDEHEVLHV"
join(complement(AL023847.1:428. .733),
complement(AL023847.1:102. .376),complement(14855. .14993))
/gene="Y57A10C.1"
/note="cDNA EST Yk247g9.5 comes from this gene; cDNA EST
Yk247g9.3 comes from this gene"
/codon_start=1
/protein_id="CAB05753.1"
/db_xref="PID:el347403"
/db_xref="PID:g3877803"
/db_xref="GI:3877803"
/translation="MGVGRILSVIYVFLGFINISYANNITVHFGTDIAQVVPYIV
AIDPLVAVPVVLFIRRGKALPIOMFYFNQVILIASAVFLSICLPMIMTPPIKSLTT
TLAVVILHCPVPVILLCISIEVHLVSPIOYTFKTRLSNLOKFKISAFPFILLEY
YITGSRDLNHWPTNIIILITAINIMTIPALVAFYACNADSIHSSRVQSGGLIVTRKW
IMKGMVNLHDGHPDEHEDICD"


```

/db_xref="GI:3057130"
/translation="MATTYEEFAAKLDRDAEFAKKMEQNKRRFFADKPDEATLSPEN
KEHYEFKMIQHTDKFNKKMREHSEHFKAFAELLEQCKNAQFPFGK"
1298..1576
/gene="KMP-11/2"
/gene="KMP-11/2"
/notes="membrane associated protein"
/codon_start=1
/product="kinetoplastid membrane protein 2"
/protein_id="AAC38991.1"
/db_xref="PID:g3057131"
/db_xref="GI:3057131"
/translation="MATTYEEFAAKLDRDAEFAKKMEQNKRRFFADKPDEATLSPEN
KEHYEFKMIQHTDKFNKKMREHSEHFKAFAELLEQCKNAQFPFGK"
1871..2149
/gene="KMP-11/3"
/gene="KMP-11/3"
/notes="membrane associated protein"
/codon_start=1
/product="kinetoplastid membrane protein 3"
/protein_id="AAC38992.1"
/db_xref="PID:g3057132"
/db_xref="PID:g3057132"
/db_xref="GI:3057132"
/translation="MATTYEEFAAKLDRDAEFAKKMEQNKRRFFADKPDEATLSPEN
KEHYEFKMIQHTDKFNKKMREHSEHFKAFAELLEQCKNAQFPFGK"
2464..2742
/gene="KMP-11/4"
/gene="KMP-11/4"
/notes="membrane associated protein"
/codon_start=1
/product="kinetoplastid membrane protein 4"
/protein_id="AAC38993.1"
/db_xref="PID:g3057133"
/db_xref="PID:g3057133"
/db_xref="GI:3057133"
/translation="MATTYEEFAAKLDRDAEFAKKMEQNKRRFFADKPDEATLSPEN
KEHYEFKMIQHTDKFNKKMREHSEHFKAFAELLEQCKNAQFPFGK"

BASE COUNT      866 a  766 c  737 g  1006 t
ORIGIN

Query Match      72.0%; Score 14.4; DB 37; Length 3375;
Best Local Similarity 68.4%; Pred. No. 2.4e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 gcdatrtgncraanaacnc 19
||:|||||:|||||:|
Db 3212 GCTATATGACCAACACAC 3194

RESULT 13
D84200 LOCUS      D84200      801 bp      DNA      BCT      06-APR-1999
DEFINITION      Escherichia coli DNA for phosphomethylpyrimidine kinase, complete
ACCESSION      D84200
NID      g4589381
VERSION      D84200.1 GI:4589381
KEYWORDS      thid; phosphomethylpyrimidine kinase.
SOURCE      Escherichia coli
ORGANISM      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE      1 (sites)
AUTHORS      Mizote, T., Tsuda, M., Smith, D.D., Nakayama, H. and Nakazawa, T.
TITLE      Cloning and characterization of the thid/J gene of Escherichia coli
encoding a thiamin-synthesizing bifunctional enzyme,
hydroxymethylpyrimidine kinase/phosphomethylpyrimidine kinase
JOURNAL      Microbiology 145 (Pt 2), 495-501 (1999)
MEDLINE      99173753
REFERENCE      2 (bases 1 to 801)

```

```

AUTHORS      Mizote, T.
TITLE      Direct Submission
JOURNAL      Submitted (23-MAR-1996) to the DDBJ/EMBL/GenBank databases. Tomoko
Mizote, Yamaguchi Prefectural University, Department of Food &
Nutrition, 3-2-1, Sakurabatake, Yamaguchi, Yamaguchi 753, Japan
(E-mail:tmizote@po.cc.yamaguchi-u.ac.jp, Tel:0839-28-2489,
Fax:0839-28-2251)
FEATURES
    source      Location/Qualifiers
                1..801
                /organism="Escherichia coli"
                /strain="W3310"
                /db_xref="taxon:562"
                /clone_lib="Kohara 8F4"
                /map="46 min"
                1..801
                /gene="thid"
                1..801
                /gene="thid"
                /codon_start=1
                /translation="MKRINALTAGTDPGAGIADKTFSGALGAYGCSVITALVAQ
                NTRGVSVYRIEDFVAQDLSVFSVDYRDTTKIGMLAETDIVEVAERLQYQIONV
                VLDIVMLAKSGDPLLSAVATLRSRLLPQVSLTPNLPFAALLDAPHARTQCEMLE
                QGRSLAMGCGAVLMKGGHLDQSDPWLFTREGEQFTAPRIMKKNHTGTGCTLSAA
                LAALRPRTNWADTVQAKSWLSALQAADTLEVHGIGVPHHFAWW"
                178 a  215 c  229 g  179 t
BASE COUNT      178 a  215 c  229 g  179 t
ORIGIN

Query Match      69.0%; Score 13.8; DB 1; Length 801;
Best Local Similarity 65.0%; Pred. No. 5.2e+02;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gcdatrtgncraanaacnc 20
||:|||||:|||||:|
Db 605 GCATTATGACCAAAACACC 624

RESULT 14
D90848/c LOCUS      D90848      17913 bp      DNA      BCT      21-MAR-1997
DEFINITION      E.coli genomic DNA, Kohara clone #359(46.8-47.2 min.).
ACCESSION      D90848 AB001340
NID      g1736814
VERSION      D90848.1 GI:1736814
KEYWORDS      Complete and shotgun sequencing; ISS; gata; gatB; gatC; gatY; gatZ;
yehA; yehB.
SOURCE      Escherichia coli (strain:K12) DNA, clone_lib:Kohara lambda miniset
library clone:Kohara clone #359.
ORGANISM      Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE      1 (bases 1 to 17913)
AUTHORS      Mori, H.
TITLE      Direct Submission
JOURNAL      Submitted (12-DEC-1996) to the DDBJ/EMBL/GenBank databases.
Hirotsada Mori, NARA Institute of Science and Technology, Res. &
Edu. Center for Genetic Info.: 8916-5 Takayama, Ikoma, Nara 630-01,
Japan (E-mail:hmori@etc.aist-nara.ac.jp, Tel:81-7437-2-5660,
Fax:81-7437-2-5669)
REFERENCE      2 (sites)
AUTHORS      Ikeda, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T.,
Ikemoto, K., Inada, T., Isono, K., Itoh, T., Kanai, K.,
Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M.,
Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H.,
Motomura, K., Nakamura, Y., Nashimoto, H., Nishio, Y., Oshima, T.,
Saito, N., Sampei, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C.,
Yamamoto, Y. and Yano, M.

```

TITLE JOURNAL REFERENCE AUTHORS

The systematic sequencing of the Escherichia coli genome in Japan
Unpublished (1996)

3 (sites)
Itoh,T., Aiba,H., Baba,T., Fujita,K., Hayashi,K., Inada,T.,
Isono,K., Kasai,H., Kimura,S., Kitakawa,M., Kitagawa,M., Makino,K.,
Miki,T., Mizobuchi,K., Mori,H., Mori,T., Motomura,K., Nakade,S.,
Nakamura,Y., Nashimoto,H., Nishio,Y., Oshima,T., Saito,N.,
Sampai,G., Seki,Y., Sivasubram,S., Tagami,H., Takeda,J.,
Takemoto,K., Wada,C., Yamamoto,Y. and Horiuchi,T.
A 460-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 40.1-50.0 min region on the linkage map
DNA Res. 3 (6), 379-392 (1996)

TITLE JOURNAL MEDLINE COMMENT

97251358
Collaboration Information:

Project:

The Japan E.coli genome DNA sequencing project

Group:

The Japan E.coli genome DNA sequencing group

Members: (1995.4 - 1996.3)

Aiba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A.,
Horiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S.,
Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S.,
Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K.,
Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K.,
Nakamura,Y., Nashimoto,H., Nishio,Y., Oshima,T., Saito,N.,
Sampai,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
Yamamoto,Y. and Yano,M.

Headed by:

Name: Takashi Horiuchi

Address: National Institute of Basic Biology, Okazaki, 444, Japan

E-mail: kishori@nibb.ac.jp

Information operator:

Name: Hirozada Mori

Address: NARA Institute of Science and Technology,

Ikoma, 630-01, Japan

E-mail: hmori@gtc.aist-nara.ac.jp

URL:

The Japan E. coli genome database

http://bsw3.aist-nara.ac.jp.

FEATURES source

Location/Qualifiers

1..17913

/organism="Escherichia coli"

/strain="K12"

/db_xref="taxon:562"

/clone="Kohara clone #359"

/clone_lib="Kohara lambda miniset library"

/map="46.8 min"

/note="Nucleotide position 2176551-2194463 from the
initiation site of ThrA (0 min.). This clone is from
Kohara lambda miniset library."

complement(1..1003)

/gene="gatC"

complement(<1..1003)

/gene="gatC"

/note="ORF_ID:0358#2; similar to [SwissProt Accession

Number P371891]"

/codon_start=1

/transl_table=11

/product="PTS system, Galactitol-specific IIC component

(EIIIC-GAT) (Galactitol- permease IIC component)

(Phosphotransferase enzyme II, C component)."

/protein_id="BAA15959.1"

/db_xref="PID:d1016684"

/db_xref="PID:g1736815"

/db_xref="GI:1736815"

/translation="MFSEVRYIILDGPTVMLPIVILIFSKILGKADCFKAGLHIG

IGFVGIGLVGLMDSIGFAAKMAENEDNLHVVDVGNFSSPMWASQALVAIPI

AILVNVALLFMTRVNDIWNIMHTFTGALLHLATGSMWIMGAGVVIHAAVYIKL

GDWFARDTRNFEELEGIAPHTGAYMGPIAVLDVDAIEKIPGVNRIKFSADDIQRF

GPFGPVTGVGMGLIIGILAGYDKVGLQAVTAAVMLMPRIKPIPMDDLPIAK

QARSRLQAFGGQEFLIGDPAALLLGHATVAVSASLIFPLTILIAVCVPGNQVLPFGD

"

complement(1007..1291)

gene

CDS

/gene="gatB"
complement(1007..1291)
/gene="gatB"
/note="ORF_ID:0358#3; similar to [SwissProt Accession
Number P37188]"
/codon_start=1
/transl_table=11
/product="PTS system, Galactitol-specific IIB component
(EIIIB-GAT) (Galactitol- permease IIB component)
(Phosphotransferase enzyme II, B component) (EC
2.7.1.69)."
/protein_id="BAA15960.1"
/db_xref="PID:d1016685"
/db_xref="PID:g1736816"
/db_xref="GI:1736816"
/translation="MKRKIIIVACGAVATSTMAAEIKELCONHNIPVELIQCRVNEI
EYMDGVHLICTAKVDRSFGDIPLVHGMFFSIGIGIEALQNKILTILOG"
complement(1322..12973)
/gene="gatA"
complement(1322..1645)
/gene="gatA"
/note="ORF_ID:0358#4; similar to [SwissProt Accession
Number P37187]; start codon is not identified yet"
/codon_start=1
/transl_table=11
/product="PTS system, Galactitol-specific IIA component
(EIIIA-GAT) (Galactitol- permease IIA component)
(Phosphotransferase enzyme II, A component) (EC
2.7.1.69)."
/protein_id="BAA15961.1"
/db_xref="PID:d1016686"
/db_xref="PID:g1736817"
/db_xref="GI:1736817"
/translation="LMNPNVLMLOWNTMISYOGVTRTFPKSSAIYLLRPTNKVHFOQA
DDNDVAVSLVIALIVENQQQLKLLRCLFGLKQDPDIVEITLITLPETQLKEYFTKYV
LSDSE"
1605..2621
/gene="IS5"
1605..2621
/gene="IS5"
/note="ORF_ID:0359#1; similar to [PIR Accession Number
A91483]"
/codon_start=1
/transl_table=11
/protein_id="BAA15962.1"
/db_xref="PID:d1016687"
/db_xref="PID:g1736818"
/db_xref="GI:1736818"
/translation="MFVINSRHTGFTMSHOLTFADEFSKRSQTRKEIFLSRMEQIL
PWQNVVEIEPFYPKAGNGRRPIPLETMRHICMQHYNISDGMEDALYEIASMRLE
ARLSLDSALPDRTTIMNFRHLEQHLAQLFKTIINRWLAEGVMMTQGLVDTATIE
APSSTNKEQRPDEMHTKKGQWFGKHAHIGVDAKSGLTHSLVTTAAEHDLNOL
GNLLHGEQFVSADAGYQGAQRELAEDVDVWLIARPGKVRTLKQHPKNTAINI
EYNKASIRARVEHFERLIIRKQFGEVKARYKGLKNDNLQNLAMLFTLANLFRADQMIRQW
ERSH"
complement(1917..2273)
/gene="IS5"
/note="ORF_ID:0359#2; similar to [PIR Accession Number
B91483]"
/codon_start=1
/transl_table=11
/protein_id="BAA15963.1"
/db_xref="PID:d1016688"
/db_xref="PID:g1736819"
/db_xref="GI:1736819"
/translation="MQQITQLIEVLMVGRGDQAVGQATIGIDTNVGLHAKVPLIAFL
GLMHLRTALLFLVLRAGCLNDGIDQALSHHDACFGQPAIDGLEQLAGLMLQQV
AEIHDGAVRQGAIQG"
complement(2659..2973)
/gene="gatA"
/note="ORF_ID:0359#4; similar to [SwissProt Accession
Number P37187]"
/codon_start=1

gene

CDS

gene

CDS

CDS

CDS

Fri Oct 1 16:47:14 1999

```

/transl_table=11
/product=PTS system, Galactitol-specific IIA component
(EIIA-GAT) (galactitol- permease IIA component)
(Phosphotransferase enzyme II, A component) (EC
2.7.1.69) .
/protein_id="BAAL5364.1"
/db_xref="PID:d1016689"
/db_xref="PID:g1736820"
/db_xref="GI:1736840"
/translation="MTNLFVRSGISFVDRSEVLTIGHNEMLAKVHVDTFPQALIARE
AFTFTGMLQEQHAIAIPHEGAEAHKAGKNGNSRLTQSFHFMEFPISPYNALNQ
pi"
complement(2983..4245)
gene

```

```

gene
complement(2983. .4245)
/gene="gatZ"
complement(2983. .4245)
/gene="gatZ"
/note="ORF ID:O359#5; similar to [SwissProt Accession
Number P37191]"
/codon_start=1
/transl_table=11
/product="Putative tagatose 6-phosphate kinase GatZ (EC
2.7.1.-)".
/protein_id="BAA15965.1"
/db_xref="PID:d1016690"
/db_xref="PID:g1736821"
/db_xref="GI:1736821"

```

```

Query Match          69.0%; Score 13.8; DB 1; Length 17913;
Best Local Similarity 85.0%; Pred. No. 5.e-02;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      1  gcdatrtgncraanacnc 20
      ||| : ||| ||| ||| |||
Db 11833 GCATTATGACCAAAACACC 11814

```

RESULT	15
HECUREASE	HECUREASE
LOCUS	2664 bp DNA BCT 03-MAY-1995
DEFINITION	Helicobacter heilmannii urease, complete cds's.
ACCESSION	L25079
NID	G793908
VERSION	L25079.1 GI:793908
KEYWORDS	urease.
SOURCE	Helicobacter heilmannii (individual_isolate 2) DNA.
ORGANISM	Helicobacter heilmannii Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group; Helicobacter. 1 (bases 1 to 2664) Solinick J.V., O'Rourke, J., Lee, A. and Tompkins, L.S. Molecular analysis of urease genes from a newly identified uncultured species of Helicobacter Infect. Immun. 62 (5), 1631-1638 (1994) 94222523 On May 3, 1995 this sequence version replaced gi:529422. Location/Qualifiers 1..2664 source
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
COMMENT	
FEATURES	

```

i. .2004
/organism="Helicobacter heilmannii"
/isolate="2"
/db_xref="taxon:35817"
/lab_host="BALB/c"
200. .204
211. .915
/codon_start=1
/transl_table=11
/product="urease"
/protein_id="AAA65722.1"
/db_xref="PID:G529423"
/db_xref="GI:529423"
/translation="MKLTPRELKMLHAGELAKORAKGIKLNYPVALISAHVM
EARAGKASVADLMQEGRTLLKDDVMPGVAMHIEVGIEAGFDGIKLVLTHTTPVEA
GSKDLKAVENEDITLANGRHAVQLVKYKNGRPVQGVSHFFEVNKLKLDFORE
KAYKRLDIASGTNRPEEGEKTVELIDIGGNKRTYGFNALVDROADDGCKLALKR

```

```

AKEKHFGTINCQCDNK "
919 . 923
930 ..2636
/codon_start=1
/transl_table=11
/product="urease"
/protein_id="AAA65723.1"
/db_xref="PID:G529424"
/db_xref="GI:529424"
/translation="MKKISRKEYSMVGYPTTIDVDTGIVKADIGKNGKHGICGAQNK
EGGGKTRDGMQTSPPSSHELDIVTNALVDYTGIVKADIGKNGKHGICGAQNK
DLQDGVNCLVGPATEALAAEGLIVTAGGIDTHIFISPOQIPTAFASGITTMIGGG
TQPDGGINATTTIPGRNWLKEMLRASEAYANMILCYLKGWVSEPALIDPLEAGAIGF
KHEDMGTSPSAINHALNIAKIDNQVAIHDTLINEAGCVEDILEIATAGTITHTPTE
GAGGAGHAPDVIKMGAEINLLPASTNPTPTFKNTEAAEHMDLMCMCHLDLKNKEDVFA
DSRIRPOTAAEDKLDHMGIFRISITSSSQAMRGVEVITWTQADKNKKEFGRLPEE
KGDNDNFRIKYISRYTINPAITHGISYGVSEVGYADLVMSPAFFGKGNMILIK
GFNIALQSMDGANASITPQQPVYRMEFGHGKAKFTNITFVSQAYANGIKHEJGL
QRVLDPVKNCRNITRKDLKFNEDVTAHIEVNPETYKKVKVDGNEVNTSHAADKLSLAQLYN
LF"
799 a * 709 c 590 g 566 t
BASE COUNT
ORIGIN
Query Match 69.0%; Score 13.8; DB 1; Length 2664;
Best Local Similarity 65.0%; Pred. No. 5.4e+02;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

Qy 1 gdatrtgncraanacnc 20
||:|:|:| ||:|:| ||:|:|
Db 2137 GCTATATTTCCAAATACACC 2156

Search completed: October 1, 1999, 15:34:56
Job time: 6199 sec

us-09-026-400-6.rge

Fri Oct 1 16:47:14 1999

PA (SUMO) SUMITOMO CHEM CO LTD.
PI Mori S, Nakanishi H, Takahashi M;
DR WPI; 98-439341/38.
PT New nicotianamine aminotransferase protein and DNA - useful for
PT enhancing iron absorption of plant cells
PS Example 2; Page 7; 17pp; English.
CC The primers V48149 and V48150 were used in the cloning of nicotianamine
CC aminotransferase can be used in a plasmid to transform plant cells to
CC produce cells with enhanced iron absorption, and it is implied [though
CC not stated] that plants with improved resistance to iron deficiency
CC chlorosis in calcareous soils can be regenerated from the transformed
CC cells. The gene fragment can be used to detect, amplify and/or isolate
CC nicotianamine aminotransferase genes. 2 G; 2 T;
SQ Sequence 20 BP; 4 A; 6 C; 2 G; 2 T;

Query Match 77.0%; Score 15.4; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gcdatrtgncraaancnc 20
|||||
DB 1 GCDATRTGNCRAAANCNC 20

RESULT 3
V48148/c
ID V48148 standard; cDNA to mRNA; 1910 BP.
AC V48148;
DT 27-OCT-1998 (first entry)
DE Nicotianamine aminotransferase 58148.62 molecular weight protein, gene.
KW ds; nicotianamine aminotransferase; plant; iron absorption;
KW iron deficiency chlorosis.
OS Gramineae sp.
FH Key Location/Qualifiers
FT 76. .1731
FT CDS /*tag= a
FT /product= "Nicotianamine transferase"
PN EP-860499-A2.
PD 26-AUG-1998. 102891.
PF 19-FEB-1998; JP-037499.
PR (SUMO) SUMITOMO CHEM CO LTD.
PI Mori S, Nakanishi H, Takahashi M;
DR WPI; 98-439341/38.
DR P-PSDB; W61643.
PT New nicotianamine aminotransferase protein and DNA - useful for
PT enhancing iron absorption of plant cells
PS Claim 4; Page 14-15; 17pp; English.
CC The nicotianamine aminotransferase can be used in a plasmid to transform
CC plant cells to produce cells with enhanced iron absorption, and it is
CC implied [though not stated] that plants with improved resistance to iron
CC deficiency chlorosis in calcareous soils can be regenerated from the
CC transformed cells. The gene fragment can be used to detect, amplify
CC and/or isolate nicotianamine aminotransferase genes. 534 C;
SQ Sequence 1910 BP; 462 A; 546 G; 368 T;

Query Match 69.0%; Score 13.8; DB 1; Length 1910;
Best Local Similarity 65.0%; Pred. No. 61;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 gcdatrtgncraaancnc 20
|||||
DB 1181 GTGATGTGCCCAACACTCC 1162

RESULT 4
V03325
ID V03325 standard; cDNA; 1253 BP.
AC V03325.
DT 15-APR-1998 (first entry)
DE Sequence of clone 15, which encodes a TRAF2 binding protein.

KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;
KW intracellular signalling activity; acute hepatitis;
KW autoimmunity-induced cell death; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 2. .1252
FT CDS /*tag= a
FT /note= "no ATG start or STOP codon given"
PN W09737016-A1.
PD 09-OCT-1997.
PF 01-APR-1997; IL0117.
PR 26-AUG-1996; IL-119133.
PR 02-APR-1996; IL-117800.
PA (YEDA) YEDA RES & DEV CO LTD.
PI Boldin M, Kovalenko A, Malinin N, Mett I, Wallach D;
DR WPI; 97-503101/46.
DR P-PSDB; W42401.
PT DNA encoding tumour necrosis factor receptor-associated factor
PT binding molecule - used for modulation or mediation in cells of the
PT activity of NF-kB
PS Claim 4; Fig 5; 127pp; English.
CC The present sequence is that of clone 15, which encodes a TRAF2
CC binding protein. Clone 15 is a partial clone, which lacks most of
CC its 5' end of the coding DNA sequence. A cDNA library prepared from
CC B-cells was screened for proteins that associate with TRAF2, and
CC the present sequence isolated. The clone 15 protein is capable of
CC binding to at least amino acids 222-501 of TRAF2. The TRAF-2 binding
CC proteins can be used for modulation or mediation in cells of the
CC activity of NF-kappaB or any other intracellular signalling activity
CC modulated or mediated by TRAF2. TRAF-binding proteins are especially used
CC for prevention or treatment of pathological conditions associated with
CC NF-kB induction, e.g. acute hepatitis, autoimmunity-induced cell death,
CC e.g. death of the beta Langerhans cells or the pancreas that results in
CC diabetes, the death of cells in graft rejection, the death of
CC oligodendrocytes in the brain in multiple sclerosis, and AIDS-inhibited
CC the AIDS disease. The proteins are also useful for screening of ligands
CC capable of binding to a protein, which are useful for modulating
CC cellular activity modulated/mediated by TRAF2.
SQ Sequence 1253 BP; 341 A; 321 C; 334 G; 256 T;

Query Match 67.0%; Score 13.4; DB 1; Length 1253;
Best Local Similarity 70.6%; Pred. No. 96;
Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 gcdatrtgncraaanc 17
|||||
DB 207 GCTATGTGCCCGAACAC 223

RESULT 5
V21210/c
ID V21210 standard; DNA; 58407 BP.
AC V21210;
DT 10-NOV-1998 (first entry)
DE Methanococcus jannaschii large circular extrachromosomal element.
DE Methanococcus jannaschii; methanogenic archaea; circular chromosome;
KW genome; autotrophic; extrachromosomal element; identification; ds.
OS Methanococcus jannaschii.
PN W09807830-A2.
PD 26-FEB-1998. U14900.
PF 22-AUG-1997; US-024428.
PR (GENO-) INST GENOMIC RES.
PA (UNIV) UNIV ILLINOIS FOUND.
PA (UIJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI Buit CJ, Smith HO, Venter JC, White OR, Woese CR;
DR WPI; 98-169145/15.
PT Complete genome sequence of methano-genic archaeon, Methanococcus
PT jannaschii - useful in identification of M. jannaschii genome
PT fragment

PS Claim 13: Page 585-600; 614pp; English.
 CC The present sequence represents the large circular extrachromosomal
 CC element sequence of the *Methanococcus jannaschii* circular chromosome. The
 CC present invention describes *M. jannaschii* open reading frames from the
 CC genome sequence. The invention also describes a computer based system
 CC for identifying fragments of the *M. jannaschii* genome that are
 CC homologous to target nucleotide sequences, comprising: (a) data storage
 CC means comprising the nucleotide sequence of the 1664976, 58407 or 16550
 CC bp sequence (see V21209, V21210 and V21211), or a nucleotide sequence at
 CC least 99.9% identical to it; (b) search means for comparing a target
 CC sequence to the nucleotide sequence of the data storage means to
 CC identify a homologous sequence, and (c) retrieval means for obtaining
 CC the homologous sequence. The method, which is based on whole genome
 CC random sequencing of an autotrophic archaeon *M. jannaschii*, the genome
 CC of which consists of 3 physically distinct elements, a large circular
 CC chromosome (the 1664976 bp sequence given in V21209), a large circular
 CC extra-chromosomal element (the 58407 bp sequence given in V21210), and a
 CC small circular extra-chromosomal element (the 16550 bp sequence given in
 CC V21211), can be used in the identification of *M. jannaschii* genome
 CC fragment.
 SQ Sequence 58407 BP; 20163 A; 9034 C; 7447 G; 21761 T;

Query Match 67.0%; Score 13.4; DB 1; Length 58407;
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;
 Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 cdattgtgccraanaacc 19
 :||:| ||:| ||:| ||:|
 Db 51951 CAATATGCCCAATATAC 51934

RESULT 6

V74830
 ID V74830 standard; DNA; 3122 BP.
 AC V74830;
 DT 16-MAR-1999 (first entry)
 DE Staphylococcus aureus contig SEQ ID #519.
 KW Computer readable medium; vaccine; S. aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 OS Staphylococcus aureus.
 FH Key Location/Qualifiers
 FT misc_feature 1021..1080
 FT /*tag= a
 FT /note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 2821..2880
 FT /*tag= b
 FT /note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT EP-786519-A2.
 PD 30-JUL-1997.
 PF 07-JAN-1997; 100117.
 PR 05-JAN-1996; US-009861.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC; Choi GH, Dillion PJ, Fannon MR, Kunsch CA.
 PI Rosen CA;
 DR WPI; 97-374922/35.
 PT Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -
 PT stored on computer readable medium and used in the production of
 PT anti-S. aureus vaccines
 PS Claim 1; Page 1446-1448; 3271pp; English.
 CC This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S. aureus DNA sequences allows putative functions to be assigned so

CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against *S. aureus* infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S. aureus in a sample. S. aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S. aureus DNA sequences contained on the
 CC computer readable medium.
 SQ Sequence 3122 BP; 1117 A; 400 C; 585 G; 899 T;

Query Match 67.0%; Score 13.4; DB 1; Length 3122;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 datrtgncraanaacc 20
 :||:| ||:| ||:| ||:|
 Db 31 AATGTGTCCCAATATACCC 48

RESULT 7

Q67186/c
 ID Q67186 standard; DNA; 41 BP.
 AC Q67186;
 DT 22-MAR-1995 (first entry)
 DE Primer for amplifying rev responsive element of HIV-1 HXB2 strain.
 KW Immunisation; vaccine; therapy; prophylaxis; defective gene;
 KW non-functional gene; template; antisense; ribozyme; bupivacaine;
 KW HIV; human immunodeficiency virus; ss.
 OS Synthetic.
 PN WO9416737-A.
 PD 04-AUG-1994.
 PF 26-JAN-1994; U00899.
 PR 26-JAN-1993; US-008342.
 PR 11-MAR-1993; US-029336.
 PR 15-JUL-1993; US-093235.
 PR 21-SEP-1993; US-124962.
 PR 21-SEP-1993; US-125012.
 PA (CONE/) CONEY L R.
 PA (MERY/) MERYA M J.
 PA (WANG/) WANG B.
 PA (WEIN/) WEINER D B.
 PA (WILL/) WILLIAMS W V.
 PA (ZURA/) ZURAWSKI V R.
 PI Coney LR, Merya MJ, Wang B, Weiner DB, Williams WV;
 PI Zurawski VR;
 DR WPI; 94-263787/32.
 PT Method for introducing genetic material into cells - utilises
 PT polynucleotide function enhancer and nucleic acid free of
 PT retroviral particles, e.g. HIV immunisation
 PT Example 49; Page 110; 136pp; English.
 CC Genetic material may be introduced into the cells of an individual
 CC by (a) contacting the individual's cells with a polynucleotide
 CC function enhancer (bupivacaine) and (b) administering to the cells
 CC the nucleic acid molecule free of retroviral particles. Nucleic
 CC acid molecules which are delivered to cells may serve as genetic
 CC templates for proteins that function as prophylactic and/or
 CC therapeutic immunising agents; replacement copies of defective,
 CC missing or non-functional genes; genetic templates for therapeutic
 CC proteins; genetic templates for antisense molecules or as genetic
 CC templates for ribozymes. Two primers (Q67186, Q67187) were used
 CC to amplify the rev responsive element (RRE) from a plasmid construct
 CC comprising the RRE sequence from HIV-1 strain HXB2. The amplified
 CC sequence was used in the construct designated pGAGPOL.rev2, used to
 CC express HIV gag and pol genes.
 SQ Sequence 41 BP; 10 A; 2 C; 17 G; 12 T;

Query Match 66.0%; Score 13.2; DB 1; Length 41;
Best Local Similarity 70.6%; Pred. No. 76;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 attgncraaacncc 20
DB 17 ATATGCCCAAAACCCC 1

RESULT 8
Q67187/c
ID Q67187 standard; DNA; 70 BP.
AC Q67187;
DT 22-MAR-1995 (first entry)
DE Primer for amplifying rev responsive element of HIV-1 HXB2 strain.
KW Immunisation; vaccine; therapy; prophylaxis; defective gene;
KW non-functional gene; template; antisense; ribozyme; bupivacaine;
KW HIV; human immunodeficiency virus; ss.
OS Synthetic.
PN W09416737-A.
PD 04-JAN-1994.
PF 26-JAN-1994; U00899.
PR 26-JAN-1993; US-008342.
PR 11-MAR-1993; US-029336.
PR 15-JUL-1993; US-093235.
PR 21-SEP-1993; US-124962.
PR 21-SEP-1993; US-125012.
PA (CONE/) CONEY L R.
PA (MERY/) MERY M J.
PA (WANG/) WANG B.
PA (WEIN/) WEINER D B.
PA (WILL/) WILLIAMS W V.
PA (ZURA/) ZURAWSKI V R.
PI Coney LR, Merva MJ, Wang B, Weiner DB, Williams WV;
PI Zurawski VR;
DR WPI; 94-263787/32.
PT Method for introducing genetic material into cells - utilises
PT polynucleotide function enhancer and nucleic acid free of
PT retroviral particles, e.g. HIV immunisation
PS Example 49; Page 110; 136pp; English.
CC Genetic material may be introduced into the cells of an individual
CC by (a) contacting the individual's cells with a polynucleotide
CC function enhancer (bupivacaine) and (b) administering to the cells
CC the nucleic acid molecule free of retroviral particles. Nucleic
CC acid molecules which are delivered to cells may serve as genetic
CC templates for proteins that function as prophylactic and/or
CC therapeutic immunising agents; replacement copies of defective,
CC missing or non-functional genes; genetic templates for therapeutic
CC proteins; genetic templates for antisense molecules or as genetic
CC templates for ribozymes. Two primers (Q67186, Q67187) were used
CC to amplify the rev responsive element (RRE) from a plasmid construct
CC comprising the RRE sequence from HIV-1 strain HXB2. The amplified
CC sequence was used in the construct designated pgAGPOL.rev2, used to
CC express HIV gag and pol genes. 8 C; 22 G; 25 T;
SQ Sequence 70 BP; 15 A; 22 G; 25 T;

Query Match 66.0%; Score 13.2; DB 1; Length 70;
Best Local Similarity 70.6%; Pred. No. 82;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 attgncraaacncc 20
DB 47 ATATGCCCAAAACCCC 31

RESULT 9
T00870/c
ID T00870 standard; DNA; 41 BP.
AC T00870;
DT 16-MAY-1996 (first entry)
DE HIV strain HXB2 rev response element 5' primer.
KW Immunisation; disease; pathogen; genetic vaccine facilitator; saponin;

anionic lipid; lectin; oestrogen; alkyl; dimethylsulphoxide; urea; PCR;
retroviral particle; retrovirus; HIV; SIV; epitope; primer;
KW amplification; ss.
PN Synthetic.
PS W09526718-A1.
PD 12-OCT-1995.
PF 30-MAR-1995; U04071.
PR 01-APR-1994; US-221579.
PA (APOL-) APOLLON INC.
PI Carrano RA;
PI WPI; 95-358434/46.
DR Introducing genetic material into cells of an individual - by
PT contacting the cells with a genetic vaccine facilitator and a
PT nucleic acid molecule
PS Example 19; Page 50; 114pp; English.
CC Immunisation of an individual against a disease or pathogen comprises
CC introducing genetic material (a genetic vaccine) into the cell of the
CC individual by contacting the cell with a genetic vaccine facilitator
CC (GVF) selected from anionic lipids, saponins, lectins, oestrogenic cpds.,
CC (GVF) selected from anionic lipids, saponins, lectins, oestrogenic cpds.,
CC hydroxylated lower alkyls, dimethylsulphoxide (DMSO) or urea, and a
CC nucleic acid that is free of retroviral particles. The primers T00830-71
CC are used in the construction of the genetic vaccines based on sequences of
CC HIV, SIV or pathogenic bacterial epitopes.
CC The primers T00870-1 were used to amplify the rev response element
CC from HIV strain HXB2 for construction of the genetic vaccine plasmid
CC pgAGPOL.rev2. The plasmid contains the rev gene from HIV strain HX3B
CC and the gag open reading frame (ORF), part of the pol ORF and the rev
CC response element from HIV strain HXB2.
SQ Sequence 41 BP; 10 A; 2 C; 17 G; 12 T;

Query Match 66.0%; Score 13.2; DB 1; Length 41;
Best Local Similarity 70.6%; Pred. No. 76;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 attgncraaacncc 20
DB 17 ATATGCCCAAAACCCC 1

RESULT 10
T00871/c
ID T00871 standard; DNA; 70 BP.
AC T00871;
DT 16-MAY-1996 (first entry)
DE HIV strain HXB2 rev response element 3' primer.
KW Immunisation; disease; pathogen; genetic vaccine facilitator; saponin;
KW anionic lipid; lectin; oestrogen; alkyl; dimethylsulphoxide; urea; PCR;
KW retroviral particle; retrovirus; HIV; SIV; epitope; primer;
KW amplification; ss.
OS Synthetic.
PN W09526718-A1.
PD 12-OCT-1995.
PF 30-MAR-1995; U04071.
PR 01-APR-1994; US-221579.
PA (APOL-) APOLLON INC.
PI Carrano RA;
PI WPI; 95-358434/46.
DR Introducing genetic material into cells of an individual - by
PT contacting the cells with a genetic vaccine facilitator and a
PT nucleic acid molecule
PS Example 19; Page 50; 114pp; English.
CC Immunisation of an individual against a disease or pathogen comprises
CC introducing genetic material (a genetic vaccine) into the cell of the
CC individual by contacting the cell with a genetic vaccine facilitator
CC (GVF) selected from anionic lipids, saponins, lectins, oestrogenic cpds.,
CC (GVF) selected from anionic lipids, saponins, lectins, oestrogenic cpds.,
CC hydroxylated lower alkyls, dimethylsulphoxide (DMSO) or urea, and a
CC nucleic acid that is free of retroviral particles. The primers T00830-71
CC are used in the construction of the genetic vaccines based on sequences of
CC HIV, SIV or pathogenic bacterial epitopes.
CC The primers T00870-1 were used to amplify the rev response element
CC from HIV strain HXB2 for construction of the genetic vaccine plasmid
CC pgAGPOL.rev2. The plasmid contains the rev gene from HIV strain HX3B

CC and the gag open reading frame (ORF), part of the pol ORF and the rev
 CC response element from HIV strain HXB2. 15 A; 8 C; 22 G; 25 T;
 SQ Sequence 70 BP; 15 A; 8 C; 22 G; 25 T;

Query Match 66.0%; Score 13.2; DB 1; Length 70;
 Best Local Similarity 70.6%; Pred. No. 82;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 atrtgnccraanacnc 20
 ||:||||:||||:||||
 Db 47 ATATGCCCAAAACCCC 31

RESULT 11

ID T32926/c
 AC T32926 standard; DNA; 41 BP.
 DT 12-FEB-1997 (first entry)
 DE HIV-1 HXB2 rev response element (RRE) gene, PCR primer.
 KW Primer; polymerase chain reaction; PCR; amplification; HIV-1;
 KW human immunodeficiency virus type 1; HXB2; gag; pol; rev;
 KW expression plasmid; induction; mucosal immunity; RRE;
 KW response element; ss.
 OS Synthetic.
 PN WO9618390-A1.
 PD 20-JUN-1996.
 PF 15-DEC-1995; U16206.
 PR 16-DEC-1994; US-357398.
 PA (UYPE-) UNIV PENNSYLVANIA.
 PI Ugen KE, Wang B, Weiner DB;
 DR WPI: 96-300372/30.
 PT Introduction of genetic material into mucosal tissue cells - by
 PT topical or lavage admin. to induce mucosal immunity or treat disease
 PS Example 14; Page 79; 91pp; English.
 CC The primer pair T32926/27 was used for the PCR amplification of the
 CC human immunodeficiency virus type 1 (HIV-1) HXB2 rev response
 CC element (RRE) gene, to produce an amplification prod. which can be
 CC used in the prepn. of a plasmid for the expression of the HIV-1
 CC HXB2 gag and pol genes. The plasmid can be introduced into an
 CC individual's cells by topical or lavage admin. to the rectal,
 CC vaginal, urethral, sublingual or buccal mucosal tissue, to induce
 CC mucosal immunity against the expressed protein and therefore
 CC HIV-1.
 SQ Sequence 41 BP; 10 A; 2 C; 17 G; 12 T;

Query Match 66.0%; Score 13.2; DB 1; Length 41;
 Best Local Similarity 70.6%; Pred. No. 76;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 atrtgnccraanacnc 20
 ||:||||:||||:||||
 Db 17 ATATGCCCAAAACCCC 1

RESULT 12

ID T32927/c
 AC T32927 standard; DNA; 70 BP.
 DT 12-FEB-1997 (first entry)
 DE HIV-1 HXB2 rev response element (RRE) gene, PCR primer.
 KW Primer; polymerase chain reaction; PCR; amplification; HIV-1;
 KW human immunodeficiency virus type 1; HXB2; gag; pol; rev;
 KW expression plasmid; induction; mucosal immunity; RRE;
 KW response element; ss.
 OS Synthetic.
 PN WO9618390-A1.
 PD 20-JUN-1996.
 PF 15-DEC-1995; U16206.
 PR 16-DEC-1994; US-357398.
 PA (UYPE-) UNIV PENNSYLVANIA.
 PI Ugen KE, Wang B, Weiner DB;

DR WPI: 96-300372/30.
 PT Introduction of genetic material into mucosal tissue cells - by
 PT topical or lavage admin. to induce mucosal immunity or treat disease
 PS Example 14; Page 79; 91pp; English.
 CC The primer pair T32926/27 was used for the PCR amplification of the
 CC human immunodeficiency virus type 1 (HIV-1) HXB2 rev response
 CC element (RRE) gene, to produce an amplification prod. which can be
 CC used in the prepn. of a plasmid for the expression of the HIV-1
 CC HXB2 gag and pol genes. The plasmid can be introduced into an
 CC individual's cells by topical or lavage admin. to the rectal,
 CC vaginal, urethral, sublingual or buccal mucosal tissue, to induce
 CC mucosal immunity against the expressed protein and therefore
 CC HIV-1.
 SQ Sequence 70 BP; 15 A; 8 C; 22 G; 25 T;

Query Match 66.0%; Score 13.2; DB 1; Length 70;
 Best Local Similarity 70.6%; Pred. No. 82;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 atrtgnccraanacnc 20
 ||:||||:||||:||||
 Db 47 ATATGCCCAAAACCCC 31

RESULT 13

Q60144/c
 ID Q60144 standard; DNA; 360 BP.
 AC Q60144;
 DT 16-MAR-1994 (first entry)
 DE Human brain Expressed Sequence Tag EST02132.
 KW Gene transcription product; genetic markers; tagging; in vivo;
 KW transcriptions; mapping; locations; chromosomes; chromosomal; ss.
 OS Homo sapiens.
 PN WO9316178-A.
 PD 19-AUG-1993.
 PR 12-FEB-1993; U01294.
 PR 12-FEB-1992; US-837195.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD, Moreno RF, Venter CJ,
 DR WPI: 93-272882/34.
 PT Enriched oligonucleotides and corresp. sequences - used as
 PT markers for human genes transcribed in-vivo, facilitate tagging
 PT of most human genes
 PS Example 4; Page 288; 500pp; English.
 CC The Expressed Sequence Tag was isolated from a human brain cDNA
 CC library as part of a large set of ESTs which can be used as markers
 CC tagging of most human genes, for mapping locations of expressed genes
 CC on chromosomes, for individual or forensic identification, for mapping
 CC locations of disease-associated genes, for identification of tissues.
 CC type, and for prepn. of antisense sequences, probes and constructs.
 CC EST02132 has a "poor" coding probability as evaluated using the
 CC coding-region prediction program CRM. See also Q59041-Q61440.
 SQ Sequence 360 BP; 89 A; 77 C; 98 G; 93 T;

Query Match 64.0%; Score 12.8; DB 1; Length 360;
 Best Local Similarity 63.2%; Pred. No. 1.8e+02;
 Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 cdattgncraanacnc 20
 |:|:|:|:|:|:|:|:|
 Db 115 CTATGTGTCGAGAGACCCC 97

RESULT 14

Q70727
 ID Q70727 standard; cDNA; 2018 BP.
 AC Q70727;
 DT 23-MAR-1995 (first entry)
 DE TATA-binding protein-associated factor dTAFLI60 cDNA.
 KW TATA-binding protein associated factor; dTAFLI60; ss; screening;

Fri Oct 1 16:47:14 1999

KW diagnostic; therapeutic; gene transcription regulation.

OS Drosophila.

FH Key Location/Qualifiers

FT cds 70..1842

FT /*tag= a

FT W09417087-A.

PN 04-AUG-1994.

PD 28-JAN-1994; U011114.

PF 28-JAN-1993; US-013412.

PR 30-JUN-1993; US-087119.

PR (REGC) UNIV CALIFORNIA.

PA Comal L, Dynlacht BD, Hoey T, Ruppert S, Tanese N;

PI Tjian R, Wang E, Weinzierl ROJ;

DR WPI; 94-264019/32.

DR P-PSDB; R56489.

PT TATA-binding protein associated protein factors - and

PT corresponding nucleotide sequence and deriv. antibodies, useful

PT in screening, diagnostics and therapeutics

PS Disclosure; Page 79-85; 180pp; English.

CC The TATA-binding protein associated factor dRAF160 (including

CC specific antibodies and fusion products) are used in drug screening,

CC diagnostics and therapeutics. They are used in the development of

CC specific biochemical assays for screening compounds that agonise or

CC antagonise selected transcription factors involved in regulating

CC gene expression associated with human pathology.

SQ Sequence 2018 BP; 507 A; 583 C; 525 G; 403 T;

Query Match

64.0%; Score 12.8; DB 1; Length 2018;

Best Local Similarity 63.2%; Pred. No. 2.3e+02;

Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 cdatrtgncraanacncc 20

I: ||| ||: || || ||

Db 477 CACTGTGCCCGAAACCC 495

RESULT 15

T02438

ID T02438 standard; cDNA; 771 BP.

AC T02438;

DT 19-APR-1996 (first entry)

DE cDNA encoding stanniocalcin from Corpuscles of Stannius.

KW stanniocalcin; Corpuscles of Stannius; hypocalcaemia; hypercalcaemia;

KW electrolyte disorder; osteoporosis; Paget's disease; treatment; ss.

OS Homo sapiens.

FH Key

FT cds Location/Qualifiers

FT 16..759

FT /*tag= a

FT /product= prepro-stanniocalcin

FT 115..757

FT /*tag= b

FT /product= mature stanniocalcin

FT 16..114

FT /*tag= c

FT /note= "encodes prepro region"

FT W09524411-A1.

PD 14-SEP-1995.

PF 09-MAY-1994; U05136.

PR 08-MAR-1994; US-208005.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Adams MD, Olsen H;

DR WPI; 95-328227/42.

DR P-PSDB; R84522.

PT Human corpuscles of Stannius polypeptide(s) - used to treat

PT hypercalcaemia, hypocalcaemia and other electrolyte disorders

PS Claim 1; Fig 1; 41pp; English.

CC The cDNA encodes a stanniocalcin, a Corpuscles of Stannius polypeptide.

CC stanniocalcin functions as a hypocalcaemic agent, and can be used for

CC the treatment of e.g. electrolyte disorders which lead to renal, bone

CC and heart diseases, hypertension, hypercalcaemia and disorders due to

CC elevated bone resorption, e.g. osteoporosis and Paget's disease.

SQ Sequence 771 BP; 212 A; 214 C; 192 G; 153 T;

Query Match 64.0%; Score 12.8; DB 1; Length 771;

Best Local Similarity 63.2%; Pred. No. 2e+02;

Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 cdatrtgncraanacncc 20

I: ||: ||| ||: || || ||

Db 616 CACTGTGCCCGAAACACACC 634

Search completed: October 1, 1999, 15:36:21

Job time: 6119 sec

Fri Oct 1 16:47:14 1999

us-09-026-400-6.rng

Page 7

